

STIC-Biotech/Ch mLib

73941

From: Rao, Manjunath N.
Sent: Thursday, August 22, 2002 4:49 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/870113

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10C 01
Phone: 306-5681

Date: 8-22-02

Please search the following as soon as possible for application with serial number **09/870,113**

SEQ ID NO: 1 against all commercial nucleic acid databases including issued patents database and pending application database and provide a **print of all results**.

SEQ ID NO: 2, 4, 8, 10, 12 against all commercial protein databases including issued patents database and pending application database and provide a **print of all results**.

If y u have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Biot chnology Patent Examiner
Art Unit 1652, Room 10A11
Crystal Mall 1, USPTO.

RECEIVED
AUG 22 2002
STIC

Searcher: D. Schreiber
Phone: 308-4292
Location: CM1 6A03
Date Picked Up: 8/26
Date Completed: 8/29
Searcher Prep/Review: 10
Clerical: _____
Online time: 7

TYPE OF SEARCH:
NA Sequences: 1
AA Sequences: 5
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: CompuGen
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

35

QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSLSDVIHPGNSHIANGAAGCVA 180
DB 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSLSDVIHPGNSHIANGAAGCVA 180
QY 181 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTOLTMNVFQA 240
DB 181 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTOLTMNVFQA 240
QY 241 IHFTYEFLOEHFNPRYNSPSSHVLSGACAGAVAAATPLDVCCKTLNTOESIALNSH 300
DB 241 IHFTYEFLOEHFNPRYNSPSSHVLSGACAGAVAAATPLDVCCKTLNTOESIALNSH 300
QY 301 ITGHTGMSAFRTYQVGGVTAFFRGVQARVIYQIPSTAIWSVYEFKYLITKRQEW 360
DB 301 ITGHTGMSAFRTYQVGGVTAFFRGVQARVIYQIPSTAIWSVYEFKYLITKRQEW 360
QY 361 RAGK 364
DB 361 RAGK 364

RESULT 2
Q920G8
ID Q920G8 PRELIMINARY; PRT; 338 AA.
AC Q920G8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Q.-Z., Ruan Q.-G., Eckenrode S., Shi J.-D., Cruz P., Wang C.-Y.,
RA She J.-X.;
RT "A new gene which is highly expressed in NOD mice spleen."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF288621; AAL23859.1; -;
SQ SEQUENCE 338 AA; 37510 MW; BB35B1F70C56A3FE CRC64;

Query Match 60.6%; Score 1176; DB 11; Length 338;
Best Local Similarity 61.9%; Pred. No. 1.4e-93;
Matches 224; Conservative 48; Mismatches 62; Indels 28; Gaps 4;
QY 1 MELEGRAGGAGVAGGPAAGPGRSPGESALLDGLQVGRGAGGAGACRPPVQDDPSG 60
DB 1 MELRRGGVGNAAAG-----RRMDGDCRDG---GGSKDAGS----- 33
QY 61 PDYBALPAGATVTTMVGAVAGILEHCVMYPIDCVKTRMOSLOPDPAARYNVLEALWR 120
DB 34 EDYENLPTSASVTHMTAGAMAGILEHSIMYPVDSVKTRMOSLSPDKARYSYIYALKR 93
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSLSDVIHPGNSHIANGAAGCVA 180
DB 94 IMHTEGFWRPLRGLNVMVMGAGPAHAMYFACYENKRTLNDVFSHQNSHLANGVAGSMA 153
QY 181 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTOLTMNVFQA 240
DB 154 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTOLTMNVFQA 240
QY 241 IHFTYEFLOEHFNPRYNSPSSHVLSGACAGAVAAATPLDVCCKTLNTOESIALNSH 300
DB 241 IHFTYEFLOEHFNPRYNSPSSHVLSGACAGAVAAATPLDVCCKTLNTOESIALNSH 300
QY 301 ITGHTGMSAFRTYQVGGVTAFFRGVQARVIYQIPSTAIWSVYEFKYLITKRQEW 360
DB 274 NVSGRLSGMANAFRTYQVGNLGLA YFKGIQARVIYQIPSTAIWSVYEFKYLITKRQEW 333
QY 360 WR 361

DB 334 NR 335
RESULT 3
Q9NYZ2
ID Q9NYZ2 PRELIMINARY; PRT; 347 AA.
AC Q9NYZ2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HT015 PROTEIN.
GN HT015.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human hypothalamus."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF223466; AAF64141.1; -;
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWNL2.
SQ SEQUENCE 347 AA; 37828 MW; 700DE61B230E001E CRC64;
Query Match 52.2%; Score 1013; DB 4; Length 347;
Best Local Similarity 60.8%; Pred. No. 1.8e-79;
Matches 197; Conservative 41; Mismatches 58; Indels 28; Gaps 5;
QY 1 MELEGRAGGAGVAGGPAAGPGRSPGESALLDGLQVGRGAGGAGACRPPVQDDPSG 60
DB 1 MELR---SGSV---GSQAVARRMDGDS-----RDGGGG-----KDATGS 33
QY 61 PDYBALPAGATVTTMVGAVAGILEHCVMYPIDCVKTRMOSLOPDPAARYNVLEALWR 120
DB 34 EDYENLPTSASVTHMTAGAMAGILEHSIMYPVDSVKTRMOSLSPDKARYSYIYALKR 93
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSLSDVIHPGNSHIANGAAGCVA 180
DB 94 IMRTEGFWRPLRGLNVMVMGAGPAHAMYFACYENKRTLNDVFSHQNSHLANGVAGSMA 153
QY 181 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTOLTMNVFQA 240
DB 154 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTOLTMNVFQA 240
QY 241 IHFTYEFLOEHFNPRYNSPSSHVLSGACAGAVAAATPLDVCCKTLNTOESIALNSH 300
DB 241 IHFTYEFLOEHFNPRYNSPSSHVLSGACAGAVAAATPLDVCCKTLNTOESIALNSH 300
QY 300 HITGHITGMSAFRTYQVGGVTA 323
DB 274 NISGRSLSGMANAFRTYQVGNLGLA 297
RESULT 4
Q969G8
ID Q969G8 PRELIMINARY; PRT; 177 AA.
AC Q969G8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE
DE CARRIER SPLICING VARIANT).
GN HMR3/4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=21195335; PubMed=11297739;
 RA Li F., Nikali K., Grehan J., Leibiger I., Leibiger B., Schweyen R.,
 RA Larsson C., Suomalainen A.;
 RT "Characterization of a novel human putative mitochondrial transporter
 homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.";
 RL FEBS Lett. 494:79-84(2001).
 DR EMBL: AJ303078; CAC27997.1; -;
 DR EMBL: AF327403; AAK49520.1; -;
 SQ SEQUENCE 177 AA; 20045 MW; E5370283AAF732E3 CRC64;

Query Match 48.3%; Score 937; DB 4; Length 177;
 Best Local Similarity 100.0%; Pred. No. 2.7e-73;
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 MNPAAVVKORMQVNSPYHRVTDVAVWQNEGAGAFYSYTTQLTWNVPFQAIHFMTYE 247
 DB 1 MNPAAVVKORMQVNSPYHRVTDVAVWQNEGAGAFYSYTTQLTWNVPFQAIHFMTYE 60
 QY 248 FLOEHFNPQRRYNPSHVLGACAGAVAAATPLDVCKTLNTQESLALNSHITGHTG 307
 DB 61 FLOEHFNPQRRYNPSHVLGACAGAVAAATPLDVCKTLNTQESLALNSHITGHTG 120
 QY 308 MASAFRTYQVGGVTAIFRGVQARVIYQIPSTAIAMSVYEFKYLITKROEWRAK 364
 DB 121 MASAFRTYQVGGVTAIFRGVQARVIYQIPSTAIAMSVYEFKYLITKROEWRAK 177

RESULT 5
 Q9H2J3 PRELIMINARY; PRT; 176 AA.
 AC Q9H2J3;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE NP0016.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PITUITARY;
 RA Xu X., Yang Y., Gao G., Xiao H., Chen Z., Han Z.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF267854; AAG44723.1; -;
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 2.
 SQ SEQUENCE 176 AA; 19931 MW; 4A4831AC572309D4 CRC64;

Query Match 46.9%; Score 910; DB 4; Length 176;
 Best Local Similarity 100.0%; Pred. No. 5.9e-71;
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VVKORMQVNSPYHRVTDVAVWQNEGAGAFYSYTTQLTWNVPFQAIHFMTYEFLOEH 252
 DB 5 VVKORMQVNSPYHRVTDVAVWQNEGAGAFYSYTTQLTWNVPFQAIHFMTYEFLOEH 64
 QY 253 FNPQRRYNPSHVLGACAGAVAAATPLDVCKTLNTQESLALNSHITGHTGMSAF 312
 DB 65 FNPQRRYNPSHVLGACAGAVAAATPLDVCKTLNTQESLALNSHITGHTGMSAF 124
 QY 313 RTYQVGGVTAIFRGVQARVIYQIPSTAIAMSVYEFKYLITKROEWRAK 364
 DB 125 RTYQVGGVTAIFRGVQARVIYQIPSTAIAMSVYEFKYLITKROEWRAK 176

RESULT 6
 Q23125 PRELIMINARY; PRT; 312 AA.
 ID Q23125
 AC Q23125;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE W02B12.9 PROTEIN.
 GN W02B12.9.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Swinburne J., Ainscough R.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z66521; CAA91399.1; -;
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
 SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;

Query Match 41.2%; Score 799; DB 5; Length 312;
 Best Local Similarity 49.5%; Pred. No. 5e-61;
 Matches 159; Conservative 54; Mismatches 82; Indels 26; Gaps 5;

QY 42 GGGEAGACGPPVRCDDPSGPDYALPAGATVTHWAGAVAGILCHCVMPIDCVKTRMQ 101
 DB 3 GGGE-----DEYESLPT-HSVPVHLTAGAGAVEHCVMFPFDSVKTRMQ 46
 QY 102 SLOPDPAARYRNVLALRIIRTEGLWRMRGLNVTATGAGPAHALYFACYEKLKLTSD 161
 DB 47 SLPCPCPETKCTPVHSLMSIVKREGVRLPLRGVNAAGSMPAHALYFTVYKMKGYLTG 106
 QY 162 VIHPGGNSH-IANGACCVATLLHDAAMNPAEVVVKORMQVNSPYHRVTCVRAVWQNEG 220
 DB 107 --NSAGHSNTLAYGAGVATLIHDAIMNPAEVVVKORMQVNSPYGSSLECARCVNREG 164
 QY 221 AGAFYSYTTQLTWNVPFQAIHFMTYEFLOEHFNPQRRYNPSHVLGACAGAVAAAT 280
 DB 165 VAAFYSYTTQLTWNVPFQAIHFMSYEFWQHVLPNPEHKYDPKSHLIAGGLAGLAAALT 224
 QY 281 PLDVCKTLNTQESLALNS-----HITGHITGMSAFRTYQVGGVTAIFRGVQARVI 333
 DB 225 PMDCVKTVLNTQAAEADPANRRIFLQARYRYRGISDAVRTIYSQRLSGFSGCGLQARVI 284
 QY 334 YQIPSTAIAMSVYEFKYLIT 354
 DB 285 FQVPATALSWSVLELKFMLS 305

RESULT 7
 Q9VAY3 PRELIMINARY; PRT; 379 AA.
 ID Q9VAY3;
 AC Q9VAY3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CG4963 PROTEIN (GH09840P).
 GN CG4963.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutcliffe G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottker P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Wellsteinbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y, CN BW SP;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC -!- INNER MEMBRANE (BY SIMILARITY).
 CC EMBL; AE003763; AAF56764.1; -;
 DR EMBL; AY060268; AAL25307.1; -;
 DR FlyBase; FBgn0039561; CG4963.
 DR InterPro; IPR001993; Mitochondrion.
 DR InterPro; IPR002067; Mitochondrion.
 DR Pfam; PF00153; mito_carrier.
 DR PRINTS; PR00926; MITOCHARRIER.
 DR PROSITE; PS00215; MITOCHARRIER; 2.
 KW Inner membrane; Mitochondrion; Transmembrane; Transport.
 SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 40.3%; Score 781; DB 5; Length 379;
 Best Local Similarity 52.7%; Pred. No. 2.3e-59;
 Matches 154; Conservative 41; Mismatches 83; Indels 14; Gaps 4;

Qy 62 DYEALPAGATVTHMVAGAVAGILEHCVMYPTDCVKTQMSLQDPDPAARVNLVLEALWRI 121
 Db 5 DYESLPT-TSVGVNMTAGIAGVLEHVVMYPLDSVKTQMSL--SPPTKNNIYVTLRTM 61
 Qy 122 IRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLLKKTLSDVHPGNSHIANGAGCVAT 181

Db 62 ITREGLRPIRGASAVVLGAGPAHSLYFAAYEMTKELAKFTSVRNLVVISGA---VAT 118
 Qy 182 LLHDAAMPNPAEVVKORMOMYNSPHYRVTDCAVAVQWQNSGAGAFYRSYTTOLTMMVFPQAI 241
 Db 119 LHDAISSPTDVIKORMOMYNSPYTSVVRDIIYKRGFRAYRAYGTQLVNLPLPYOTI 178
 Qy 242 HFMTYEFQEHFNPPORRYNSPSHVLGACAGAAAAATTPDLVDCKTLLTNTQSLALNSHI 301
 Db 179 HFTTYEFFQNMNLKERNPPVHMAAGAAAGACAAVTPDLVDIKTLNTQE----- 230
 Qy 302 TGHITGASAFRTVYQGVGTAYFRGVQARVYIQIPSTAIAMSVYEFFKYLI 353
 Db 231 TGLTRGMIEASRKIYHMAGLPLGFRGTTRVLYSMPATAICWSTYEFFKFYL 282
 RESULT 8
 Q94638 ID PRELIMINARY; PRT; 303 AA.
 AC Q94638;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE MITOCHONDRIAL SOLUTE CARRIER.
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OX NCBI_TaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96326580; PubMed=8703971;
 RA Catnull J., Miller D.J.;
 RT "cDNAs from *Onchocerca* sp. encoding members of the MRS3/MRS4 class of
 RT mitochondrial solute carriers.";
 RL Biochim. Biophys. Acta 1282:179-181(1996).
 DR EMBL; U45998; AAB19037.1; -;
 DR InterPro; IPR001993; Mitochondrion.
 DR Pfam; PF00153; mito_carrier.
 DR PROSITE; PS00215; MITOCHARRIER; UNKNOWN_2.
 SQ SEQUENCE 303 AA; 33861 MW; 224BFA547E5D617C CRC64;

Query Match 38.3%; Score 742.5; DB 5; Length 303;
 Best Local Similarity 45.2%; Pred. No. 3.7e-56;
 Matches 142; Conservative 54; Mismatches 87; Indels 31; Gaps 3;

Qy 49 CRPPVRODPSGPDYALPAGATVTHMVAGAVAGILEHCVMYPTDCVKTQMSLQDPDPA 108
 Db 12 CRWPV-----HLLAGSVAGLAELHCLMFPDPSVKTQLQSLCPCE 50
 Qy 109 ARYRNLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLLKKTLSDVHPGNS 168
 Db 51 TSCPAMHLSMVKREGLLSLKGVNAVLTIPAHAFYTVVYNSKAYLLN--NPRVS 108
 Qy 169 SHIANGAAGCVATLLHDAAMPNPAEVVKORMOMYNSPHYRVTDCAVAVQWQNSGAGAFYRSY 228
 Db 109 NSVSAISGALATVTHDAVMNPAEVVKORMOMYNSPHYRVTDCAVAVQWQNSGAGAFYRSY 168
 Qy 229 TTQTMNVFPQAIHMTYEFQEHFNPPORRYNSPSHVLGACAGAAAAATTPDLVDCKTL 288
 Db 169 ITQTLNVPIYOCYTHPMIYEQNLNPHHDYNPSSHVSGGTAGGIAAATTPDLCVKTV 228
 Qy 289 LNTQES-----LALNSHITGHITGMASAFRTVYQGVGTAYFRGVQARVYIQIPSTA 340
 Db 229 LNTQTPRENTYRLLTQSEHTAYKGLADGKTIYLRGTGGFGRGLQARIIPSTA 288
 Qy 341 IAWSVYEFFKYLIIT 354
 Db 289 LSWSAVELCKYMLS 302
 RESULT 9
 Q9P0J2 ID PRELIMINARY; PRT; 187 AA.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 02:34:08 ; Search time 138.54 Seconds
(without alignments)
291.835 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGVAGGPAAGPG.....VYEFKYLTKRQEWRAAGK 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1696	87.4	318	22 AAM41505	Human polypeptide
2	1534	79.1	289	21 AAB50383	Human uncoupling p
3	1534	79.1	289	22 AAB60113	Human transport pr
4	1387	71.5	268	22 AAM39719	Human polypeptide
5	1113.5	57.4	366	22 AAM00938	Human bone marrow
6	1025.5	52.9	272	21 AAB42980	Human ORFX ORF2744
7	1012.5	52.2	331	22 AAB60658	Human mitochondria
8	919	47.4	677	22 AAU29748	Novel human secret
9	916	47.2	174	22 AAU01989	Human secreted pro
10	910	46.9	176	21 AAB40634	Human ORFX ORF398
11	910	46.9	176	22 AAB98221	Human mitochondria

12	856.5	44.1	299	22 AAM79039	Human protein SEQ
13	781	40.3	379	22 ABB61130	Drosophila melanog
14	739.5	38.1	199	22 AAM80023	Human protein SEQ
15	725.5	37.4	187	22 AAM93909	Human polypeptide,
16	725.5	37.4	187	22 AAB98222	Human mitochondria
17	725.5	37.4	187	22 AAB49665	Human polypeptide,
18	717.5	37.0	187	22 AAM93934	Human secreted pro
19	717.5	37.0	188	21 AAB32112	O. volvulus mitoch
20	558	28.8	215	22 AAB49666	Human secreted pro
21	541	27.9	194	21 AAB32111	Gene #25 human sec
22	515.5	26.6	181	22 AAU01988	Blackcurrant Rib7
23	510	26.3	289	18 AAM17054	Arabidopsis thalia
24	509	26.2	331	21 AAG22077	Arabidopsis thalia
25	507	26.1	331	21 AAG43092	Arabidopsis thalia
26	497	25.6	290	21 AAG22078	Arabidopsis thalia
27	495	25.5	290	21 AAG43093	Arabidopsis thalia
28	453	23.4	278	21 AAG22079	Arabidopsis thalia
29	451	23.2	278	21 AAG43094	Arabidopsis thalia
30	440.5	22.7	155	21 AAB42966	Human ORFX ORF2730
31	412	21.2	111	21 AAB42968	Human ORFX ORF2732
32	343.5	17.7	149	21 AAB42964	Human ORFX ORF2728
33	309	15.9	74	22 ABB01167	Novel human diago
34	308.5	15.9	413	22 ABB69108	Drosophila melanog
35	305.5	15.7	449	22 ABB59263	Drosophila melanog
36	300	15.5	301	21 AAG52781	Arabidopsis thalia
37	300	15.5	301	21 AAG52783	Arabidopsis thalia
38	300	15.5	325	21 AAG52780	Arabidopsis thalia
39	300	15.5	325	21 AAG52782	Arabidopsis thalia
40	300	15.5	361	21 AAG52779	Arabidopsis thalia
41	294.5	15.2	358	22 ABB71142	Drosophila melanog
42	293	15.1	351	21 AAB50378	Human uncoupling p
43	281.5	14.5	374	22 AAM40959	Human polypeptide
44	281.5	14.5	374	22 AAM40960	Human polypeptide
45	281	14.5	365	21 AAB42783	Human ORFX ORF2547

ALIGNMENTS

RESULT 1

AAM41505
ID AAM41505 standard; Protein; 318 AA.

XX AC AAM41505;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 6436.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI60661.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 2; SEQ ID NO 6436; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAI38642-AAI62213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-brager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 318 AA;
 SQ

Query Match 87.4%; Score 1696; DB 22; Length 318;
 Best Local Similarity 100.0%; Pred. No. 1.6e-163;
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GACPPVRQDPDGPDEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMOSLOPD 106
 DB 1 gacppvrqdpdgpdealpagaatvtthmvagavagilehcvmypidcvktrmqslqpd 60
 QY 107 PAARYRNVLALWRIIRTEGLRPMRGLNVTATGAGPAHALYFACYEKLKTKTSDVIHPG 166
 DB 61 paaryrnvlalwriirteglrpmrglnvtatgagpahalyfacyeklkktlsdvihpq 120
 QY 167 GNSHIANGAGCVATLLHDAAMPVAVKORMQYNPSYHRVTDCVRAVWQNEGAGAFYR 226
 DB 121 gnschiangagcvatllhdaampvavkqrmqynpsyhrvtdcvravwqnegagafyr 180
 QY 227 SYTTOLTMNVYFQAIHPTMYEFLQEHFNPPORRYNPSSHVLSCACAGAVAAATPLDVCK 286
 DB 181 syttqlcmnvypfqaiahptmyeflqehfnpprrynpsshvlsqacagavaaatpldvck 240
 QY 287 TLLNTQESLALNSHITGHITGMSAFRTVYQVGGVTAAYFRGQVARYIQTIPSTAIWSVY 346
 DB 241 tllntqeslalnshithitgmsaftrvyqvggvtaayfrgqvaryiqtipstaiawsy 300
 QY 347 EFPKYLITKQEWRAK 364
 DB 301 effkylitkrqeewragk 318

RESULT 2
 AAB50383
 ID AAB50383 standard; Protein; 289 AA.
 XX
 AC AAB50383;
 DT 12-MAR-2001 (first entry)
 XX
 DE Human uncoupling protein #6.
 XX

Human; uncoupling protein; immunosuppressive; antiarthritic;
 antiinflammatory; antiproliferative; cardiant; vasotropic;
 cerebroprotective; neuroprotective, antibacterial; ophthalmological;
 gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;
 gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
 infertility.
 Homo sapiens.
 WO2000061614-A2.
 19-OCT-2000.
 06-APR-2000; 2000WO-US09534.
 09-APR-1999; 99US-0128701.
 08-JUL-1999; 99US-0142821.
 18-AUG-1999; 99US-0149448.
 12-NOV-1999; 99US-0164751.
 (HUMA-) HUMAN GENOME SCI INC.
 Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;
 WPI; 2000-656322/63.
 N-PSDB; AAC90457.
 Uncoupling proteins and nucleic acid sequences encoding them, useful
 for detecting, preventing and treating proliferative, neurological,
 immune system, cardiovascular and gastrointestinal disorders -
 Claim 11; Page 323-324; 343pp; English.
 The present sequence is a human uncoupling protein. The nucleotide
 sequences encoding the uncoupling proteins may be used for
 the detection of various disorders such as cancer, for chromosome
 identification, as chromosome markers and for numerous other diagnostic
 or research purposes. The uncoupling protein encoded by the nucleotide
 sequences may be used to treat disorders such as neural, immune,
 muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 renal and proliferative disorders, wounds, infectious diseases,
 thrombosis, arthritis, and infertility.
 Sequence 289 AA;
 Query Match 79.1%; Score 1534; DB 21; Length 289;
 Best Local Similarity 100.0%; Pred. No. 3.8e-147;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 MVAGAVAGILEHCVMYPIDCVKTRMOSLOPDPAARYRNVLALWRIIRTEGLRPMRGLN 135
 DB 1 mvagavagilehcvmypidcvktrmqslqpdpaaryrnvlalwriirteglrpmrgln 60
 QY 136 VTATGAGPAHALYFACYEKLKTKTSDVIHPGNSHANGAGCVATLLHDAAMPVAVK 195
 DB 61 vtatgagpahalyfacyeklkktlsdvihpqngshiangagcvatllhdaampaevvk 120
 QY 196 QRMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQTLTMNVYFQAIHPTMYEFLQEHFNP 255
 DB 121 qrmqynpsyhrvtdcvravwqnegagafyrsyttqtltmnvypfqaiahptmyeflqehfnp 180
 QY 256 QRRYNPSSHVLSCACAGAVAAATPLDVCKTLLNTQESLALNSHITGHITGMSAFRTV 315
 DB 181 qrrynpsshvlsqacagavaaatpldvcktllnqeslalnshithitgmsaftrv 240
 QY 316 YQVGGVTAAYFRGQVARYIQTIPSTAIWSYEFFKYLITKQEWRAK 364
 DB 241 yqvggvtaayfrgqvaryiqtipstaiawsyeffkylitkrqeewragk 289

RESULT 3
 AAB60113

ID AAB60113 standard; Protein; 289 AA.
AC AAB60113;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33.
XX
KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO200078953-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16668.
XX
PR 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
DR WPI; 2001-041424/05.
DR N-PSDB; AAF27733.
XX
PT Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
PS Claim 2; Page 133-134; 165pp; English.
XX
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated tppts). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
XX Sequence 289 AA;

Query Match 79.1%; Score 1534; DB 22; Length 289;
Best Local Similarity 100.0%; Pred. No. 3.8e-147;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 MVAGAVAGILEHCVMYPIDCVKTRMOSLOPDPAPARYNVLEALWRIIRTEGLWRPMLN 135
Db 1 mvagavagilehcvmypidcvktrmslqdpaparynvlealwriirteglwrpmln 60
QY 136 VTATGAGPAHALYFACYEKLKTLSDVIHPGGNSHTANGACCVATLLHDAAMNPAEVVK 195
Db 61 vtatgagpahalafacyeklkctlsdvihpggnshtangagcgvatllhdaamnpaevvk 120
QY 196 QRMQNTSPYHRTVDCVRWQNEGAGAFYSYTTQITMNVFOATHFMTYEFLOQHFPN 255
Db 121 qrmqntspyhrtvdcvrwqnegagafysyttltmnpvfqaihftmtyeflqehfnp 180
QY 256 QRRYNPSSHVLGACAGAVAAATTPLDVCKTLNLTQESLALNSHITGHITGMASAFRTV 315
Db 181 qrrynpsshvlsgacagavaaattpldvcktlntlqeslalnshitghitgmasafrtv 240
QY 316 YQVGGVTAYFRGQVQARVYIQISTALAWSVYEFKYLITRKQEWRAKG 364
Db 241 yqvggvtayfrgqvqarvyiqibstaiaawsveyeffkylitkrqewragk 289

RESULT 4
AAM39719
ID AAM39719 standard; Protein; 268 AA.
XX
AC AAM39719;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2864.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-052317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI58875.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 2864; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 268 AA;

Query Match 71.5%; Score 1387; DB 22; Length 268;
Best Local Similarity 98.1%; Pred. No. 2.9e-132;
Matches 262; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 98 TRMQSLQPDPAARYNVLEALWRIIRTEGLWRPMLNVTATGAGPAHALYFACYEKLKK 157

Db 2 trmhsldsqdpgayrnyvnlvliwriirteglwrpmsglntatgawpahalyfacyekikk 61

Qy 158 TLSDVIHPGNSHIANGAGCVATLLHDAAMNPAEAVKQRMQYNSPYHRTVDCVRVWQ 217

Db 62 tilsdvihpngnsihiangaagcvatlilhdaamnpaevkqrmqynsp/hrvtdeavrwwq 121

Qy 218 NEGAGAFYRSYTTQLTMMNVPFOAIHFMTVEFLQEHFNQRRYNPSSHYLSGACAGAVAAA 277

Db 122 negagafyrsyttqltmmnvpfoahfmyeflqehfnpqrrynpsshylsagacagavaaa 181

Qy 278 ATPTLDVCKTLNTQESLALNSHITGHITGMASAFRTYVQGVGTAYFRGVQARVYIYQIP 337

Db 182 atptldvcktlntqeslalnshitghitgmasafrtvyqvggtayfrgvqarvlyqip 241

Qy 338 STAIAWSYEFKYLITKROEWRAGK 364

Db 242 staiawsyeffkylitkrgeewragk 268

RESULT 5

AAM00938

ID AAM00938 standard; Protein; 366 AA.

XX

AC

XX

XX

DT

XX

XX

01-OCT-2001 (first entry)

DE Human bone marrow protein, SEQ ID NO: 414.

XX

XX

KW Human; bone marrow; antinflammatory; cytostatic; neuroprotective; antiviral; antibacterial; antifungal; anti-HIV; haemostatic; immunosuppressive; gene therapy; cytokine cell proliferation; cell differentiation modulator; immune disorder; infection; cancer; human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.

XX

OS Homo sapiens.

XX

PN WO200153453-A2.

XX

XX

PD 26-JUL-2001.

XX

PF 23-DEC-2000; 2000WO-US34960.

XX

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 30-NOV-2000; 2000US-0250583.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;

PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao Qa, Zhou P, Drmanac RT;

XX

DR WPT: 2001-488707/53.

DR N-PSDB: AAH90057.

XX

PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful for treating e.g. cancer and immune deficiency disorders -

XX

PS Claim 10; Page 504-505; 648pp; English.

XX

CC The present sequence is one of 251 novel human polypeptides encoded by a bone marrow-expressed polynucleotide. The polynucleotide and the polypeptide encoded by it are useful in the treatment of various immune deficiencies and disorders. The deficiencies and disorders may be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal infection, or may result from an autoimmune disorder, a coagulation

CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation, suppression of an inflammatory response or treatment of a nervous system disorder such as Alzheimer's disease. Detection of the presence or increased expression of the polynucleotide or the protein it encodes is useful for the diagnosis and/or prognosis of one or more types of cancer. The polynucleotide and polypeptide can be used as nutritional sources or supplements and in the screening of chemical compounds as potential drugs.

XX

SQ Sequence 366 AA;

Query Match 57.4%; Score 1113.5; DB 22; Length 366;

Best Local Similarity 69.4%; Pred. No. 2.6e-104;

Matches 202; Conservative 43; Mismatches 45; Indels 1; Gaps 1;

Qy 73 THMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNVLALWRIIRTEGLWRPWR 132

Db 74 sthmtagmagilehsvmvpvsvktrmqslspdpkadytsiygalkkimrtgfwrlr 133

Qy 133 GLNVTATGAGPAHALYFACYEKLKLTSLDVTHPGNSHIANGAGCVATLLHDAAMNPAE 192

Db 134 gvnvmimgagpahamyfacyemkrlndvfhngnshiangagsmatllhdavmnpae 193

Qy 193 VVKQRMQYNPSYHRVTDVRAVMQNEGAGAFYRSYTTQLTMMNVPFOAIHFMTVEFLQEH 252

Db 194 vvkqrlmqnysqhrsaiscirtvrtgltgafyrtynpqltmnlpfqsihfityeflqeq 253

Qy 253 FNPQRRYNPSSHVLSGACAGAVAAAATPTLDVCKTLNTQESLALN-SHITGHITGMASA 311

Db 254 vnphtynpqsghilssglagalaaatpldvcktlntqenvalnsanigrisgmana 313

Qy 312 FRTVYQGVGTAYFRGVQARVYIYQIPSTAIAWSYEFKYLITKROEWRSA 362

Db 314 frtvyqnglgyfkgldqrvlyqmpstaiswsyeffkylitkrqlenra 364

RESULT 6

AAB42980

ID AAB42980 standard; Protein; 272 AA.

XX

AC AAB42980;

XX

XX

DT 08-FEB-2001 (first entry)

XX

XX

DE Human ORFX ORF2744 polypeptide sequence SEQ ID NO:5488.

XX

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antithyroid; antianaemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive.

XX

OS Homo sapiens.

XX

PN WO200058473-A2.

XX

XX

PD 05-OCT-2000.

XX

XX

PF 31-MAR-2000; 2000WO-US08621.

XX

XX

PR 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

```

PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
PA Shinkets RA, Leach M;
PI WPI; 2000-602362/57.
XX N-PSDB; AAC77189.
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX Claim 11; Page 4662-4663; 5507pp; English.
PS AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressive;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antihypertensive; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy.
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX Sequence 272 AA;
SQ

```

```

DE Human mitochondrial solute carrier protein hMSC-o.
XX Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
KW preparation; detection.
XX Homo sapiens.
OS CN1269409-A.
PN 11-OCT-2000.
XX 17-MAR-2000; 2000CN-0114958.
PF 17-MAR-2000; 2000CN-0114958.
XX 17-MAR-2000; 2000CN-0114958.
PR (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
PA Zhang X, Gao X, Xiao H;
XX WPI; 2001-050544/07.
DR N-PSDB; AAF59920.
XX New human mitochondrion solute carrier protein and its nucleic acid -
PT Claim 4; Page 20; 21pp; Chinese.
XX The invention relates to a novel human mitochondrial solute carrier
CC protein, hMSC-o (AAF60658), and cDNA encoding it (AAF59920). hMSC-o is
CC expressed in normal human hypothalamus tissue. The invention also relates
CC to the preparation of hMSC-o proteins and nucleic acids, and the
CC detection of hMSC-o proteins and nucleic acids in a sample. The present
CC sequence represents hMSC-o.
XX Sequence 331 AA;
SQ

```

```

Query Match 52.2%; Score 1012.5; DB 22; Length 331;
Best Local Similarity 65.1%; Pred. No. 4e-94;
Matches 188; Conservative 39; Mismatches 51; Indels 11; Gaps 2;
QY 36 GVGGRGAGGAGACRPVPRQDPDSDYALPAGATVTTMVGAVAGILEHCVMPIDC 95
Db 3 gdsrdgggg-----kdatgsdyenlptsasvsthtaganaagilehsvmypvds 52
QY 96 VKTRMOSLOPDPAARYRNVLRLIIRTEGLWRPRLMGLNVTATGAGPAHALYFACYEKL 155
Db 53 vktrmqslspdkadytslygalkkrmrtcgfwrlgrgvnmimgagpahamyfacyenm 112
QY 156 KKTLSDVIPHGNSHIANGAAGCVATLLHDAAMNPAEVVKQRMQYNSPYHRVTCVRVAV 215
Db 113 krtldvfhhgqshlangiagsmatllhdavmnpaevvkqrlqynsqhrsaiscirtv 172
QY 216 WQNEGAGAFYRSYTTQLTMMNVFQAIHFMTYEFLEHFNPRQRRYNSPSSHVLSCACAGAVA 275
Db 173 wrteglgafyrsyttqtlmnpifqslhityeflgeqvnphrtyppqshliaggagala 232
QY 276 AAATPLDVKCKLLNTQESLALN-SHITCHITGMASAFRTYVQGVGVT 323
Db 233 aaatpdlvckllntqenvaislanisgrlsmanaftrvyqlnglpa 281

```

```

RESULT 8
AAU29748
ID AAU29748 standard; Protein; 677 AA.
XX AC
XX AAU29748;
XX 18-DEC-2001 (first entry)
XX DE
XX Novel human secreted protein #239.
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

```

```

Query Match 52.9%; Score 1025.5; DB 21; Length 272;
Best Local Similarity 69.4%; Pred. No. 1.5e-95;
Matches 186; Conservative 41; Mismatches 40; Indels 1; Gaps 1;
QY 96 VKTRMOSLOPDPAARYRNVLRLIIRTEGLWRPRLMGLNVTATGAGPAHALYFACYEKL 155
Db 3 lqtrmqslspdkadytslygalkkrmrtcgfwrlgrgvnmimgagpahamyfacyenm 62
QY 156 KKTLSDVIPHGNSHIANGAAGCVATLLHDAAMNPAEVVKQRMQYNSPYHRVTCVRVAV 215
Db 63 krtldvfhhgqshlangiagsmatllhdavmnpaevvkqrlqynsqhrsaiscirtv 122
QY 216 WQNEGAGAFYRSYTTQLTMMNVFQAIHFMTYEFLEHFNPRQRRYNSPSSHVLSCACAGAVA 275
Db 123 wrteglgafyrsyttqtlmnpifqslhityeflgeqvnphrtyppqshliaggagala 182
QY 276 AAATPLDVKCKLLNTQESLALN-SHITCHITGMASAFRTYVQGVGVT 334
Db 183 aaatpdlvckllntqenvaislanisgrlsmanaftrvyqlnglpa 242
QY 335 QIPSTAIASVYEFFKYLITRKQEWRA 362
Db 243 qmpstaisvyeffkylitkrqlenra 270

```

```

RESULT 7
AAB60658
ID AAB60658 standard; Protein; 331 AA.
XX AC
XX AAB60658;
XX 04-MAY-2001 (first entry)
XX

```


|||||
Db 61 ifhmfyeflqehfnpqrrynpsshvlgacagavaaaatp|dvcktl|ntqes|alnsh 120
QY 301 ITGHTGMSAPRTYVQGVGTAYFRGVQARVIYQIPSTAIWSVYEFKYLIT 354
Db 121 itghtgmasairtyvqgvgtayfrgvqarvlyqipstaiaawsyeffkylit 174
RESULT 10
AAB40634
ID AAB40634 standard; Protein; 176 AA.
XX
AC AAB40634;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF398 polypeptide sequence SEQ ID NO:796.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR N-PSDB; AAC74843.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 854; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,

CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 176 AA;
Query Match 46.9%; Score 910; DB 21; Length 176;
Best Local Similarity 100.0%; Pred. No. 4.2e-84;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 193 VVKRMQMYNSPYHRVTDVRAVWQNEGAGAFYSYTTTLTMNVPFOAIHFMFYELQEH 252
Db 5 vvkrmqmynsphyrvtdcvravwnegagafysyttltnmvpfqaihfmfyelqeh 64
QY 253 FNPORRYNPSSHVLSGACAGAVAAATP|LDVCKTL|NTQES|ALNSHT|TGHTGMSAF 312
Db 65 fnporrynpsshvlgacagavaaaatp|dvcktl|ntqes|alnsh|tghitgmasaf 124
QY 313 RTVYQGVGTAYFRGVQARVIYQIPSTAIWSVYEFKYLITKROEWRAGK 364
Db 125 rtvyqgvgtayfrgvqarvlyqipstaiaawsyeffkylitkrqewragk 176
RESULT 11
AAB98221
ID AAB98221 standard; Protein; 176 AA.
XX
AC AAB98221;
XX
DT 20-AUG-2001 (first entry)
XX
DE Human mitochondrial solute carrier (hMSC-homologue) protein SEQ:7.
XX
KW Human; hMSC; hMSC-homologue; mitochondrial solute carrier.
XX
OS Homo sapiens.
XX
PN CN1281895-A.
XX
PD 31-JAN-2001.
XX
PF 27-JUN-2000; 2000CN-0116795.
XX
PR 27-JUN-2000; 2000CN-0116795.
XX
PA (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
PI Li N, Xu X, Xiao H;
XX
DR WPI; 2001-282781/30.
DR N-PSDB; AAB22162.
XX
PT New human mitochondrial solute carrier (hMSC) protein isomer and its
PT code sequence -
XX
PS Claim 4; Page 18 (disclosure); 22pp; Chinese.
XX
CC The present invention describes a human mitochondrial solute carrier
CC protein, designated hMSC-homologue, which is expressed in human
CC hypophysis. Also described are methods for the preparation and
CC detection of hMSC-homologue protein and nucleotide sequences. The
CC present sequence represents hMSC-homologue, as given in the present
CC invention.
XX
SQ Sequence 176 AA;
Query Match 46.9%; Score 910; DB 22; Length 176;

Best Local Similarity 100.0%; Pred. NO. 4.2e-84; Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VVKQRMQYNSPYHRVTDVCRVAVQNEGAGAFYRSYTTQLTMMVPFQAHFMTYEFLOEH 252
Db 5 vvkqrmqynspyrvtcdvcravqnegagafyrsyttqltmmvpfqahfmyeflqeh 64

QY 253 FNPQRYNPSHVLSCAGAGAVAAAATTPLDVCKTLINTQESLALNSHITGHITGMASAF 312
Db 65 fnpqrynpsshvlsagacagavaaaattpldvcktlintqeslalnshitghitgmasaf 124

QY 313 RTVYQGVGTAYPRGVQARVIYQIPSTAIAMSVEFFKYLITKRQEWRAKG 364
Db 125 rtvyqvgvtayirgvgarviyqipstaiawsyeffkylitkrqeewragk 176

RESULT 12
AAM79039
ID AAM79039 standard; Protein; 299 AA.
XX AAM79039;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1701.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK52172.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 4043; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 299 AA;

QY 1 MELEGKAGGVAGGPAAGPQRSPGESALLDQWLQGVGRGAGGAGACRPPVRQDDSG 60
Db 28 melr---sgsv--gsqavarimdgds-----rdgggg-----kdatgs 60

QY 61 PDYEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNVLALWR 120
Db 61 edyenlptsaasvsthtagamagilehsvmypydvsktrmqslspopkagysvvgalkk 120

QY 121 IIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLAKTLSDVIHPGNSHIANGAGCVA 180
Db 121 imrtgfwrlprlgvnmimgagpahamyfacyenmkrtlnldvfhqgnshlangiagsma 180

QY 181 TLLHDAAMNPAEYVKORMQYNSPYHRVTDVCRVAVQNEGAGAFYRSYTTQLTMMVPFQA 240
Db 181 tllhdavmnpaevvkqrlqmynsqhrsaiscirtvrtgltgafyrsyttqltmmvpfqs 240

QY 241 IHEWTFEFLQEHFNPPQRYNPSHVLSCAGAGAVAAAATT 280
Db 241 ihitwyeflqehfnppqrynpsshvlsagacagavaaaatt 280

RESULT 13
ABB61130
ID ABB61130 standard; Protein; 379 AA.
XX ABB61130;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 10182.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL05233.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 10182; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
DR N-PSDB; RAK94871.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 4059; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 187 AA;

Query Match 37.4%; Score 725.5; DB 22; Length 187;
Best Local Similarity 71.9%; Pred. No. 2.4e-65;
Matches 133; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

Qy 179 VATLLHDAAMNPAEVYKORMQYNSPYHRVTDVCVRYWONEGAGAFYRSYTTQLTMMVPE 238
Db :||||| ||||| ||||| :||| :||| :||| ||||| ||||| :|||
1 matilhdavmpaevvqrlqymnsqhrsaiscirtvrtgigafyrsyttqtmnlpf 60

Qy 239 QAIHMTYEFLEHFNPPORRYNPSSHLVGACAGAVAAAATPLDVCCKTLINTQESALN 298
Db :||||| ||| ||| ||| :||| :||| ||||| ||||| :|||
61 qsihftyeflqeqvnphtynpqshiisgglagalaaaattpldvcktilntqenval 120

Qy 299 -SHITGHTGNASAFRIYQVGGVTATFRGVQARVIYQIPSTATAWSVYEFKYLITKRO 357
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
121 lanisgrlsgmanaftrvyqnglagvyfgkigqrviyqmpstaiswsvvefkyfltkrq 180

Qy 358 EEWRA 362
Db :||
181 lenra 185

Search completed: August 27, 2002, 03:55:57
Job time: 4909 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 02:37:18 ; Search time 48.11 Seconds
(without alignments)
184.804 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGVGAGGPAAGPG.....VYEFKYLITKROENRAGK 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	511	26.3	328	4	US-09-068-140A-15
2	510	26.3	289	4	US-09-068-140A-10
3	293	15.1	351	2	US-08-933-750C-19
4	293	15.1	351	4	US-09-234-613-19
5	265	13.7	447	4	US-09-160-119-4
6	265	13.7	674	4	US-09-160-119-2
7	259	13.4	320	2	US-08-933-750C-12
8	259	13.4	320	4	US-09-234-613-12
9	255.5	13.2	469	4	US-09-188-930-339
10	229.5	11.8	311	2	US-08-775-009-32
11	221	11.4	311	2	US-08-775-009-33
12	201.5	10.4	432	2	US-08-937-466-4
13	201.5	10.4	432	2	US-09-172-528-4
14	201.5	10.4	432	3	US-09-318-199-4
15	201.5	10.4	432	4	US-09-503-579-4
16	198	10.2	306	5	PCT-US94-09799-1
17	197.5	10.2	308	2	US-08-937-466-2
18	197.5	10.2	308	2	US-09-172-528-2
19	197.5	10.2	308	3	US-09-318-199-2
20	197.5	10.2	308	4	US-09-503-579-2
21	186	9.6	312	4	US-09-142-565-2
22	183.5	9.5	307	2	US-08-807-861A-56
23	183.5	9.5	307	3	US-09-210-681-56
24	183.5	9.5	307	3	US-08-946-719A-56
25	181	9.3	256	2	US-08-937-466-6
26	181	9.3	256	2	US-09-172-528-6
27	181	9.3	256	3	US-09-318-199-6

28 181 9.3 256 4 US-09-503-579-6
29 171.5 8.8 309 1 US-08-518-878B-51
30 171.5 8.8 309 2 US-08-807-861A-51
31 171.5 8.8 309 2 US-08-470-868A-51
32 171.5 8.8 309 3 US-09-210-681-51
33 171.5 8.8 309 3 US-08-946-719A-51
34 168.5 8.7 299 1 US-08-518-878B-56
35 168.5 8.7 299 2 US-08-470-868A-56
36 139.5 7.2 303 1 US-08-518-878B-37
37 139.5 7.2 303 1 US-08-294-522B-36
38 139.5 7.2 303 2 US-08-807-861A-37
39 139.5 7.2 303 2 US-08-470-868A-37
40 139.5 7.2 303 3 US-09-210-681-37
41 139.5 7.2 303 3 US-08-946-719A-37
42 138.5 7.1 298 3 US-08-961-871-10
43 121.5 6.3 312 4 US-09-188-930-142
44 101.5 5.2 1155 1 US-08-094-948A-29
45 101.5 5.2 1155 5 PCT-US96-09319-29

ALIGNMENTS

RESULT 1
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-15

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYN00AT01
CLONE: 724157
US-08-933-750C-19

Query Match 15.1%; Score 293; DB 2; Length 351;
Best Local Similarity 25.1%; Pred. No. 7.7e-22;
Matches 89; Conservative 50; Mismatches 140; Indels 76; Gaps 8;
QY 67 PAGATVTTTHVAGAVAGILEHCVMYPIDCVKTRMOSLOPDPA----- 109
DB 6 PAGISPLQOMVAGSGTCAVVTSLFMTPLDVVKVRLQSQPSMASLMPSSRLWSLSYTKWK 65
QY 110 -----RYRNVLEALWRIIRTEG-----LWRPMLGLNVT 137
DB 66 CLLYCNGVLEPLYLCPNGARCATWFDPTFRFTGMDAFVKIVRHEGRTLW---SGLPAT 122
QY 138 ATGACPAHALYFACYEKLLKTL-----SDVIHPGNSHIANGACCVATLLHDAAMNPA 191
DB 123 LVMTVPATAIYFTAYDQLKAFLCGRALTSDLAP-----MVAGALARLGTVTVISPL 174
QY 192 EVVKORMQMSNPYHRVTDVCRVAVQWNEGAFYRSYTTQLTMNVPFOAIHFMTEYFLOE 251
DB 175 ELMRTKLAQHVSRELGCACVTAQAQGWRSLSLWLGWPTALRDVFFSALYFNFVELYKS 234
QY 252 HFN---PQRRYNPSSHVLGACAGAVAAATPLDVKCT-----LLNTOESLALNSHITGH 304
DB 235 WLNGLRPKDQTSVGMSEFVAGGISGTVAAVLTLPEDVVKTRQVALGAMEAVRVNPL---H 291
QY 305 ITGMSAFRTVYQGVGTAYFRGVQARVIYQIPSTAIAMSVYEFKYLITKQEE 359
DB 292 VDSWTLRLRRRAESGTGKLFAGFLPRIKAAPSCAIMISTYEFKGFQRLNQD 346

RESULT 4

US-09-234-613-19
Sequence 19, Application US/09234613
Patent No. 6132973
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Nell C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/234,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYN00AT01
CLONE: 724157
US-09-234-613-19
Query Match 15.1%; Score 293; DB 4; Length 351;
Best Local Similarity 25.1%; Pred. No. 7.7e-22;
Matches 89; Conservative 50; Mismatches 140; Indels 76; Gaps 8;
QY 67 PAGATVTTTHVAGAVAGILEHCVMYPIDCVKTRMOSLOPDPA----- 109
DB 6 PAGISPLQOMVAGSGTCAVVTSLFMTPLDVVKVRLQSQPSMASLMPSSRLWSLSYTKWK 65
QY 110 -----RYRNVLEALWRIIRTEG-----LWRPMLGLNVT 137
DB 66 CLLYCNGVLEPLYLCPNGARCATWFDPTFRFTGMDAFVKIVRHEGRTLW---SGLPAT 122
QY 138 ATGACPAHALYFACYEKLLKTL-----SDVIHPGNSHIANGACCVATLLHDAAMNPA 191
DB 123 LVMTVPATAIYFTAYDQLKAFLCGRALTSDLAP-----MVAGALARLGTVTVISPL 174
QY 192 EVVKORMQMSNPYHRVTDVCRVAVQWNEGAFYRSYTTQLTMNVPFOAIHFMTEYFLOE 251
DB 175 ELMRTKLAQHVSRELGCACVTAQAQGWRSLSLWLGWPTALRDVFFSALYFNFVELYKS 234
QY 252 HFN---PQRRYNPSSHVLGACAGAVAAATPLDVKCT-----LLNTOESLALNSHITGH 304
DB 235 WLNGLRPKDQTSVGMSEFVAGGISGTVAAVLTLPEDVVKTRQVALGAMEAVRVNPL---H 291
QY 305 ITGMSAFRTVYQGVGTAYFRGVQARVIYQIPSTAIAMSVYEFKYLITKQEE 359
DB 292 VDSWTLRLRRRAESGTGKLFAGFLPRIKAAPSCAIMISTYEFKGFQRLNQD 346

RESULT 5

US-09-160-119-4
Sequence 4, Application US/09160119A
Patent No. 6316219
GENERAL INFORMATION:
APPLICANT: KRIEF, STEPHANE
APPLICANT: SOUCHET, MICHEL
APPLICANT: BRIL, ANTOINE
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GH-30985
CURRENT APPLICATION NUMBER: US/09/160,119A
CURRENT FILING DATE: 1998-09-24
EARLIER APPLICATION NUMBER: EP 97402511.6
EARLIER FILING DATE: 1997-10-23
EARLIER APPLICATION NUMBER: EP 98401655.0
EARLIER FILING DATE: 1998-07-02
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 4
; LENGTH: 447
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-4

          13.7%;   Score 265;   DB 4;   Length 447;
Best Local Similarity 28.2%;   pred. No. 8.3e-19;
Matches      82;   Conservative 47;   Mismatches 122;   Indels 40;   Gaps

QY    79  GAVAGAILHCVMYPIDCVKTRMQSLQPPA-----ARYRNVLALWRIIRTEGLWRPMRGL 134
       I::I:: :III::IIII:: : : : : : : : : : : : : : : : : : : : : : : :
Db    107 GSVAGAGATAVYPIDLVKTRMQNRSTGSGFVGMELMYKNPSDFCKFKLVRYEGFFGLYRGL 166
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY    135 NVTATGACPAHALFYACEYKLIKLTSDVI-----HPGCN-----SHIANGAAGCAVTLIH 184
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    167 LPQLLGVAPEKA-----IKLTVDNFVRDKFMHKDGSVPAAEILAGGCGAGSQVIF- 217
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY    185 DAAMNPAAEVVKORMQMYNSPHYRVDCVR----AWQNNEGAGAFYSRYTTLTMNVPOQA 240
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    218 ---TNPLEIVAIRLOVAG---EITTPGRVSALSVVRLDLGGFYKGAKACFLRDIPESA 270
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY    241 IHETMYEFPLQEHF--NQRRYNNPSSHLSVGACAGAVAATAATPLDYCKTLLNTQESLAUNS 299
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    271 IYFCYAHHVKASFANEDQSVPGSLLAGATAGMPAAPASLVTPADVIKTRLQVAARAGOTT 330
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY    300 HITGHITGMSAFRTVTVQGVGTAYFRGVQVARVIYQIPSTAIAWSVYEKK 350
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    331 Y-----SGVIDCFKILLKEEGPKALKWGAGARVRSSPFGFWTLITYELLO 376
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT      6
US-09-160-119-2
; Sequence 2, Application US/09160119A
; Patent No. 6316219
; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRIL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160.119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 674
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-160-119-2

```

	Query Match	13.78;	Score 265;	DB 4;	Length 674;
	Best Local Similarity	28.28;	Pred. No. 1.5e-18;		
	Matches 82;	Conservative 47;	Mismatches 122;	Indels 40;	Gaps 9;
QY	79	GAVAGILEHCVMYPIDCVKTRMQSLQDPDA----	ARYRNVLLEALWRIIRTEGLWRPMRGL	134	
		: : : : : : : :			
Db	334	GSVAGAGATAVIPIDLVKTRMQRSTGSEVGLMYKNSDFCKVKVLYRGFFGLYRGL	393		
QY	135	NVTATGAGPAHALYFCYEKLKKTLSVDI-----	HPGNN-----SHIANGAAGCAVATLLH	184	
		: : : : : :			
Db	394	LPQLLGVAPEKA-----	IKLVNDFVRDKFMKDGVSPLAAEILAGGCAGGSQVIF-	444	
QY	185	DAAMNPAEVYKQRMQNTNSPHYRVTDVCR----	AWONEGAGAYRSYTTOLTNNVPFQA	240	
		: : : : : :			
Db	445	---TNPLEIVRIQLQVAG----	ETTTGPRYSALSVSRDLGFFYIGKAKACFLRDIPFSA	497	

```

QY      241 IHFWTYEFLQEHF-NPQRYNPSSHVLSCAGAVAAAATTPLDCKTKLLNTQESIALNS   299
Db      498 IYFCYAHVKASFANEDGOVSPGSLLAGATAGMPAASLVTPADVIKTRLOVAARAGOTT   557
QY      300 HITGHITGMASAFRTVYGVGWTFYFRYQVARVIYQIPSTAIANSVYEFFK   350
Db      558 Y-----SGVIDCFKRILREEGPKALWKAGARVRFPSSPQFGVTLLTYELLQ   603

RESULT          7
US-08-933-750C-12
; Sequence 12, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Puri
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNNOT02
; CLONE: 207452
; US-08-933-750C-12

```

	Query Match	13.4%	Score 259;	DB 2;	Length 320;
	Best Local Similarity	24.7%;	Pred. No. 2.le-18;		
	Matches	78;	Conservative 62;	Mismatches 134;	Indels 42; Gaps 12:
Qy	63 YEALPAGATVTHM--VAGAVAGILEHCVMYPIDCVKTRMSLO-----PDPAARYRNV	114	: : : :		
Dd	4 YDPKDGNNRTKFQVAVAGSVGLVTRALISFFDFVIKIRFO-LQHERLSRSDPSAKYHGI	62	: : : :		
Qy	115 LEALWRIIRTGEG---LWRPMRGLVNTAGCAPAHLYACFYCKLUKTLDSVDIHPGG----	167	: : : :		

```
Db 63 LQASRQILOEEGPTAFKWHVPAQILSIGYGAOVFLSF-----EMLTELIVHRGSYDA 115
QY 168 ---NSH-IANGAGCVATLLHDAAMNPAEVVKQRMQYNSP--YHRYTDCVRAVWQNEGA 221
Db 116 REFSVHFVCGGLAACMATL-----TVHPVDVLRTRFAAQGEKPVYNTLRHAVGTMYRSEGP 171
QY 222 GAFYRSYTTQITMNVFQAITHFMYEFLQEHFN-----PQRRYNPSSHVLSGACAGAVAA 276
Db 172 QVFYKGLAPTLLIAFPYAGLQFSCYSSLKHLKYKWAIPAEGKKNENLQNLCCSGAGVISK 231
QY 277 AATTPLDVCKTLTNT---QESLALNSHITGHITCMASAFRTVYQVGVTAFFRGVQARVI 333
Db 232 TLTYPLDLFKRLQVGGFEHARAAGVQVRRY-KGLMDCAKQVLQKEGALGFFKGLSPSLL 290
QY 334 QYIPSTAIAMSVYEFF 349
Db 291 KAALSTGCMFFSYEFF 306

RESULT 8
US-09-234-613-12
; Sequence 12, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNOT02
; CLONE: 207452
US-09-234-613-12
```

```
Query Match 13.4%; Score 259; DB 4; Length 320;
Best Local Similarity 24.7%; Pred. No. 2.1e-18;
Matches 78; Conservative 62; Mismatches 134; Indels 42; Gaps 12;

QY 63 YEALPAGATVTHM--VAGAVAGILEHCVMYPIDCVTRMQSLQ-----PDPAARTNV 114
Db 4 YDPKPDGRNNTKFOVAVAGSVGLVTRALISPFDVIRFQ-LQHERLSRSDPSAKYHGI 62
QY 115 LEALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKLTLSDVHHPGG--- 167
Db 63 LQASRQILOEEGPTAFKWHVPAQILSIGYGAOVFLSF-----EMLTELIVHRGSYDA 115
QY 168 ---NSH-IANGAGCVATLLHDAAMNPAEVVKQRMQYNSP--YHRYTDCVRAVWQNEGA 221
Db 116 REFSVHFVCGGLAACMATL-----TVHPVDVLRTRFAAQGEKPVYNTLRHAVGTMYRSEGP 171
QY 222 GAFYRSYTTQITMNVFQAITHFMYEFLQEHFN-----PQRRYNPSSHVLSGACAGAVAA 276
Db 172 QVFYKGLAPTLLIAFPYAGLQFSCYSSLKHLKYKWAIPAEGKKNENLQNLCCSGAGVISK 231
QY 277 AATTPLDVCKTLTNT---QESLALNSHITGHITCMASAFRTVYQVGVTAFFRGVQARVI 333
Db 232 TLTYPLDLFKRLQVGGFEHARAAGVQVRRY-KGLMDCAKQVLQKEGALGFFKGLSPSLL 290
QY 334 QYIPSTAIAMSVYEFF 349
Db 291 KAALSTGCMFFSYEFF 306

RESULT 9
US-09-188-930-339
; Sequence 339, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 339
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-339
```

```
Query Match 13.2%; Score 255.5; DB 4; Length 469;
Best Local Similarity 28.6%; Pred. No. 8.4e-18;
Matches 84; Conservative 51; Mismatches 120; Indels 39; Gaps 14;

QY 75 HMVAGAVAGILEHCVMYPIDCVTRMQSLQPDPAARYN--VLEALWRIIRTEG---LWR 129
Db 189 HLVAGGAGAVSRCTAPLDRKLVMQV---HASRSNNMCIVGGFTQMIREGAKSLWR 244
QY 130 PMRGLNVTATGAGPAHALYFACYEKLKLTLSDV---IHPGGNSHTANGAAGCVATLLH 184
Db 245 G-NGINVLK--IAPESAIFKFMAYEQMKRLVGSQDETURIH---ERUVAGSLAGIA---- 294
QY 185 DAAMNPAEVVKQRMQYNS-PYHRYTDCVRAVWQNEGAFYRSYTTQITMNVFQAIFH 243
Db 295 QSSIYPMEVLKTRMALRKITGOYSGMLDCARRILAKEGVAAFYKGYIPNMLGIIPYAGIDL 354
QY 244 MYEFLQEHENPQRRY-----NPSHVL--SGACAGAVAAATTPLDVCKTLNTQESLA 296
Db 355 AVVETLKNW--LQRYAVNSADPGVFLVACGTTISSTCGQLASYPALVRTRMQAASIE 412
```



```
QY 288 LLNTOESLALNSHITGHTGMSAFRTVYQVG-----GVTAYFRGVOARVIYQIPSTA 340
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 RMQGLE-----AHKYRNTLDCGVQILKNEGPRAFYKGTVPRLGRVCLDYA 291
QY 341 IAWSYVEFKYLITK 355
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
292 IVFVIYDEVVKLLNK 306

RESULT 12
US-08-937-466-4
; Sequence 4, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-937-466-4

Query Match 10.4%; Score 201.5; DB 2; Length 432;
Best Local Similarity 25.3%; Pred. No. 2.6e-12;
Matches 78; Conservative 44; Mismatches 139; Indels 47; Gaps 12;

QY 71 TWTTHWAGAVAGILEHCVMYPIDCVKTRMOSLQDPAA---RYRNVLKLTSDVHPGNSHIANG 127
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
13 TTVKFLGAGTAACFAD-LLTPLDPTAKVRLQIQENPGQASVQYRGVLGTILTMVTEGP 71
QY 128 WRPMRGLNVTATGACPAHALYFA-----CYEKLKKTLSDVHPGNSHIANG---AAGCV 179
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 RSPYSGL-----VAGLHRQMSFASIRIGLYDSVKQFYT-----PKGADHSSVAIRILAGCT 122
QY 180 ATLLHDAAMNPAEVVKQRMQMY-----NSPYHRVTDVCRVAVWQNEGAGAFYRSYTTQL 232
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
123 TGAMAVTCAQPTDVVVKVRFQAMIRLGTGGERKRYRTMDAYRTIAREEGVRLGWKGTWPN 182
QY 233 TMNVFPQAIHFMTYEFLEQHFNPQRRY--NPSSHVLSCAGACAVAAAATTPLDVCKTLN 290
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
183 TRNAIVNCAEMWYTDIIIEKLLSHLFDTNFPCHVSAFGAGFCATVVASPDVVVKTRY- 241
QY 291 TQESLALNSHITGHTGMSAFRTVYQVGVTAYFRGVOARVIYQIPS--TAIAWSVYEF 348
```

```
Db 242 -----MNAPLGRYSPLCHMLKWAQ-BGPTAFYKGF-----VPSFLRLCANVMVF 287
QY 349 FKYLITKR 356
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
288 VTYEQLKR 295

RESULT 13
US-09-172-528-4
; Sequence 4, Application US/09172528
; Patent No. 5952469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/172,528
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-172-528-4

Query Match 10.4%; Score 201.5; DB 2; Length 432;
Best Local Similarity 25.3%; Pred. No. 2.6e-12;
Matches 78; Conservative 44; Mismatches 139; Indels 47; Gaps 12;

QY 71 TWTTHWAGAVAGILEHCVMYPIDCVKTRMOSLQDPAA---RYRNVLKLTSDVHPGNSHIANG 127
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
13 TTVKFLGAGTAACFAD-LLTPLDPTAKVRLQIQENPGQASVQYRGVLGTILTMVTEGP 71
QY 128 WRPMRGLNVTATGACPAHALYFA-----CYEKLKKTLSDVHPGNSHIANG---AAGCV 179
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 RSPYSGL-----VAGLHRQMSFASIRIGLYDSVKQFYT-----PKGADHSSVAIRILAGCT 122
QY 180 ATLLHDAAMNPAEVVKQRMQMY-----NSPYHRVTDVCRVAVWQNEGAGAFYRSYTTQL 232
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
123 TGAMAVTCAQPTDVVVKVRFQAMIRLGTGGERKRYRTMDAYRTIAREEGVRLGWKGTWPN 182
QY 233 TMNVFPQAIHFMTYEFLEQHFNPQRRY--NPSSHVLSCAGACAVAAAATTPLDVCKTLN 290
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
183 TRNAIVNCAEMWYTDIIIEKLLSHLFDTNFPCHVSAFGAGFCATVVASPDVVVKTRY- 241
```



```
QY 233 TMNVPQAIHMTYEFLOEHFNQRRY--NPSSHVLGACAGAVAAAAATPLDVCKTLN 290
| | :||: :| | : | | | | :||: | |
Db 183 TRNAIVNCAEMTYDIIKEKLLSHLFTDNFPCHFVSATGAGFCATVWASPDVVKTRY- 241
| | :||: :| | : | | | | :||: | |
QY 291 TQESLALNSHITGHITGMASAFETYVQGGVTAYFRGVOARVIYQIPS--TAIAWSVYEF 348
| | :||: :| | : | | | | :||: | |
Db 242 -----MNAPLGRYRSPHCLMKMVAQ-EGPTAFYKGF-----VPSFLRLGAWNVMVF 287
| | :||: :| | : | | | | :||: | |
QY 349 FKYLITKR 356
| | :||: :| | : | | | | :||: | |
Db 288 VTYEQLKR 295
| | :||: :| | : | | | | :||: | |
```

Search completed: August 27, 2002, 03:58:49
Job time: 4891 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 02:35:14 ; Search time 80.04 Seconds
(without alignments)
436.988 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGGAGGPAAGPG.....VYEFFKYLIRKQEWRAKG 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	799	41.2	312	2	T26089		hypothetical prote
2	601	31.0	303	2	T39149		probable RNA splic
3	598.5	28.8	304	2	S13533		mRNA splice defec
4	555.5	28.6	314	2	S51179		mRNA splice defec
5	521	26.9	781	2	A86205		hypothetical prote
6	507	26.1	331	2	T00582		probable mitochond
7	313.5	16.2	368	2	S54524		hypothetical prote
8	297	15.3	367	2	G36770		hypothetical prote
9	295	15.2	328	2	T19322		hypothetical prote
10	293	15.1	351	2	T43493		hypothetical prote
11	282.5	14.6	902	1	S54495		probable carrier p
12	274	14.1	269	2	T02090		hypothetical prote
13	274	14.0	330	2	T09362		hypothetical prote
14	272.5	14.0	307	2	S60949		probable phosphat
15	266.5	13.7	436	2	J01459		Btl protein precu
16	265.5	13.7	284	2	S45458		PET8 protein - yea
17	250.5	12.9	322	2	S57116		probable carrier p
18	250	12.9	338	2	T40968		probable mitochond
19	249.5	12.9	384	1	S44092		probable carrier p
20	245.5	12.7	309	2	T48156		hypothetical prote
21	240.5	12.4	475	2	T50686		peroxisomal Ca-dep
22	240	12.4	392	2	T05350		adenylate transloc
23	239.5	12.3	352	2	T01729		mitochondrial solu
24	235	12.1	326	2	S57544		probable membrane
25	234.5	12.1	335	2	T50393		probable mitochond
26	230.5	11.9	345	2	T37576		probable mitochond
27	229.5	11.8	311	2	G01789		citrate transporte
28	228.5	11.8	336	2	T01839		hypothetical prote
29	228	11.8	373	2	S48451		probable membrane

30	226	11.6	650	2	T32897	hypothetical prote
31	225.5	11.6	366	2	S64589	hypothetical prote
32	225	11.6	335	2	S50453	hypothetical prote
33	224.5	11.6	588	2	T22688	hypothetical prote
34	224	11.5	347	2	T50990	hypothetical prote
35	223	11.5	479	2	T49871	peroxisomal Ca-dep
36	222.5	11.5	302	2	T38879	probable mitochond
37	221	11.4	296	2	B96830	hypothetical prote
38	221	11.4	311	2	A46595	tricarboxylate tra
39	220	11.3	277	2	T40033	probable mitochond
40	219	11.3	271	2	S62485	probable mitochond
41	219	11.3	312	2	H89567	protein K11H3.3 [l
42	217.5	11.2	307	2	S67662	hypothetical prote
43	217	11.2	330	2	T29640	mitochondrial carr
44	216.5	11.2	349	2	B96753	hypothetical prote
45	215	11.1	702	2	T16533	hypothetical prote

ALIGNMENTS

RESULT 1

T26089

hypothetical protein W02B12.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T26089

R:Swinsburne, J.; Ainscough, R.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z20149

A:Accession: T26089

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-312 <WIL>

A:Cross-references: EMBL:Z66521; PIDN:CAA91399.1; GSPDB:GN000020; CESP:W02B12.9

A:Experimental source: clone W02B12

C:Genetics:

A:Gene: CESP:W02B12.9

A:Map position: 2

A:introns: 18/3; 251/3; 286/3

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 41.2%; Score 799; DB 2; Length 312;

Best Local Similarity 49.5%; Pred. No. 4.7e-56;

Matches 159; Conservative 54; Mismatches 82; Indels 26; Gaps 5;

QY 42 GGGEAGACRPVVRQDPDSDGPDYEALFAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQ 101

Db 3 GGGE-----DEYSLPT-HSVPVHLTAGALAGAVEHCVMFPDVSVKTRMQ 46

QY 102 SIQPPAARYRNVLKALWRIIRTEGLWRPRLNVTATAGFAHALYFACYEKLKLTSD 161

Db 47 SLCPCPETKPTPVHSILMSIVKREGRLRGNVAVAGSMPAHALYFTVYKMKGYLTG 106

QY 162 VTHPGNSH-IANGAGCVATLLHDAAMNPAEVRKQMNSPYHRVTDVCRVAVWNEG 220

Db 107 --NSAGHSNTLAYAGSGVVATLIHDAIMPAEVRKQMNSPYHRVTDVCRVAVWNEG 164

QY 221 AGAFYRSTYTLTMNVFFQAIHFMTYEFLOEHFNQRRYNPSHVLGACAGAVAAAT 280

Db 165 VAAFYRSTYTLTMNVFFQAIHFMTYEFLOEHFNQRRYNPSHVLGACAGAVAAAT 224

QY 281 PLDVCKTLTNTQESALNS-----HTGHTGTGMASAPRTVYQGVGTAYPRGQARVI 333

Db 225 PMDCVKTVLNTQAAEADPANRRIFLQARYRYRGISDAVRTIYSORGLSGFSGCGLQARVI 284

QY 334 YQIPSTATAWSVYEFKYLIT 354

Db 285 FQVPATALSWSVYELFKFMLS 305

RESULT 2

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-328 <WIL>

A:Cross-references: EMBL:Z46787; PIDN:CAA86739.1; GSPDB:GN00021; CESP:C16C10.1

A:Experimental source: clone C16C10

C:Genetics:

A:Gene: CESP:C16C10.1

A:Map position: 3

A:Introns: 22/1; 125/2; 179/1; 298/1

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match

Best Local Similarity 15.2%; Score 295; DB 2; Length 328;

Matches 101; Conservative 48; Mismatches 146; Indels 68; Gaps 14;

QY 22 SPGESALLDGLWLGVRGAGGGERACRPP-----VRQDPDSGPDYEALPAGATVTTTH 76

DB 3 SPCE-----GKTTNCSGAPSSCVTPLDVVKIRLQQQTRP-----FPRGECFYH- 48

QY 77 VAGAVAGILEHCVMYPIDCVKTRMQSLQ-----DPAARYRNVLALWRIIRTEGLWRPMR 132

DB 49 -----NGLMEHV-----CVSEVR--KPEWYQRPNGFRGADAIKTIARHEGIRSLWS 95

QY 133 GLNVATGAGPAHALYFACYEKKTKLSDV1-----HPGNSHIANGAAGCVAT 181

DB 96 GLSPTWVWALPATVFFTYDNLVWLKKMCCRRAFSPEKWTTPDWS--AAVAGIVAR 153

QY 182 LHLDAMNPAEYVVKORMQYNSPYHRVTCVRAVWQNEGAGAFYRSYTTQLTWNVPFQAI 241

DB 154 TIATVWVSPYEMIRTKMQSKRTYHIGHLVRSSMATKGISFFLGTWPTMLRDIPFSGI 213

QY 242 HPMTEYFLEQHEPNPORYNPSSH-----VLSGACAGAVAAATPLDVCCKTLTNTQESLA 296

DB 214 YWAGDLEKT--NLQRQGDHPNFVSVFSGAAGVVASIIFHPDVIKT--NCQ---- 265

QY 297 LNSHITGHTGMASAFRTV-----YQGVGTAYFRGVQARVYIQIPSTAIWSVYEFKYL 352

DB 266 --IRIGSIDDMNKSTITTVIKDMYHSGISAFSSGLVPLRVKVSPPSCAIMISFYEFKEL 323

QY 353 ITK 355

DB 324 FOX 326

RESULT 10

T43493

hypothetical protein DKF2p434C119.1 - human

C:Species: Homo sapiens (man)

C:Date: 21-Jan-2000

C:Accession: T43493

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z22516

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-351 <AA>

A:Cross-references: EMBL:AL133584

A:Experimental source: adult testis; clone DKF2p434C119

C:Genetics:

A:Note: DKF2p434C119.1

Query Match

Best Local Similarity 15.1%; Score 293; DB 2; Length 351;

Matches 89; Conservative 50; Mismatches 140; Indels 76; Gaps 8;

QY 67 PAGATVTTTHVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPA----- 109

DB 6 PAGISPLQOMVASGTGAVVTSLEMTPLDVVKVQSQRPMSAMELMPSSRLSLSYTKWK 65

QY 110 -----RYRNVLEALWRIIRTEG---LWRPMRGLNVT 137

DB 66 CLLYCNGVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVREGTRTLW----SGLPAT 122

QY 138 ATGAGPAHALYFACYEKKTKL-----SDVHPGNSHIANGAAGCVATLLHDAAMNPA 191

DB 123 LVMTPATATYFTAYDQLKAFLCGRALTSLYAP-----MVAGALARLGTVTVISPL 174

QY 192 EVVKORMQMYNSPYHRVTCVRAVWQNEGAGAFYRSYTTQLTWNVPFQAIHFMTYEFLOE 251

DB 175 ELMRTKLQAHVSYSRELGACVRTAVAGGWRSLWLGWGPPTALDRDVPFSALYFNVELVS 234

QY 252 HFN---PQRRYNPSSHVLSCACAGAVAAATTPLDVCKT-----LLNTQESLALNSHITGH 304

DB 235 WLNLGRPKDQTSVGMSFVAGGISGTVAAVLTLPFVVKTORQVALGAMAEAVRNPL--H 291

QY 305 ITGMAFAFTVQVQGVGTAYFRGVQARVYIQIPSTAIWSVYEFKYLITKRQEE 359

DB 292 VDSWLLLRIRAESGTKGLFAGFLPRIIKAAPSCAIMISTYEFKGKSFQRLNOD 346

RESULT 11

S54495

probable carrier protein YPR021c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YP9367.01c

C:Species: Saccharomyces cerevisiae

C:Date: 10-Sep-1999

C:Accession: S54495

R:Badcock, K.; Churcher, C.M.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54059

A:Accession: S54495

A:Molecule type: DNA

A:Residues: 1-902 <BAD>

A:Cross-references: EMBL:Z49274; NID:9809585; PIDN:CAA89275.1; PID:9809586; GSPDB:GN0

A:Experimental source: strain AB972

C:Genetics:

A:Gene: MIPS:YPR021c

A:Map position: 16R

C:Superfamily: probable carrier protein YPR021c; ADP,ATP carrier protein repeat homol

C:Keywords: duplication; transmembrane protein

F:527-615/Domain: ADP,ATP carrier protein repeat homology <ACP1>

F:621-711/Domain: ADP,ATP carrier protein repeat homology <ACP2>

F:724-814/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 14.6%; Score 282.5; DB 1; Length 902;

Best Local Similarity 29.4%; Pred. No. 2.1e-14;

Matches 86; Conservative 45; Mismatches 129; Indels 33; Gaps 8;

QY 79 GAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTA 138

DB 537 GSTAGCIGATVVPIDFIFIKRMAQR--SLAQYKNSIDCLLKIIISREGIKGLYSGLGQOL 594

QY 139 TGAGPAHALYFACYEKKTKLSD-----VIHPGNSHIANGAAGCVATLLHDAAMNPAEV 193

DB 595 IGVAPEKAIKLVNDPMNRNLTDKNGKLSLEFPEIISGASAGACQVIFT-----NPLEI 647

QY 194 VKORMQMYNSPY-----HRVTCVRAVWQNEGAGAFYRSYTTQLTWNVPFQAIHFMTYEF 248

DB 648 VKIRLQV-QSDYVGENIQOANETATQIVKRLGLRGLYNGVAACLMRDVPFSAIYFFTYAH 706

QY 249 LOE---HFNP-----QRRYNPSSHVLSCACAGAVAAATTPLDVCKTLTNTQESLALNSH 300

DB 707 LKKDLDFDPNDKTKRNLKLTWELLTAGAIGAMPAAFLTTPDFVIKRIQIDPRKGETKY 766

QY 301 ITGHTGMASAFRTVQVGVGTAYFRGVQARVYIQIPSTAIWSVYEFKYLII 353

DB 767 ----NGIFHAIRTIKESFRSFFKGGGARVLRSSSPQFGFTLAAEYELFKGFI 814

RESULT 12

T20290

hypothetical protein D1046.3 - Caenorhabditis elegans


```
Db 77 RLYRGISSPMLMEAPKRAKTFACNDQYQKIFKNLENTNETTKISIAAGASAGTAAVI 136
      | |::: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 189 NPAEVVYKQRMQNTSPYHRVTDVRAVWQNEGAGAYRSYTTQLTMTNVPQAIFH-MTYE 247
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 VPFLIKIRMQDVKSYLGPMDLCKTKIKNEGIMGLYKGIESTMWRNALWNGGYFGYIYQ 196
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 248 FLQEHENPORRYNP-----SSHLVGACAGAVAAATTPLDVCKTLTNTQESIALN 298
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 197 V-----RNSMPVAKTKQKTRNDLIAGIGTGTMLNTPFDVVKSRIQSV--AVS 246
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 299 SHITGHTGMASAFRTVYQVGGYAYFRGQARVYIQIPSTAIWSYI-----EFFFYL 352
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 247 SAVKKNWCLPSLL-VYIREGFRALYKGFVPKVCRLAPGSLMLVYFTGMNFFRDL 303
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 15

JQ1459

Bt1 protein precursor - maize

C:Species: Zea mays (maize)

C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 24-Sep-1999

C:Accession: JQ1459

R:Sullivan, T.D.; Strelow, L.I.; Illingworth, C.A.; Phillips, R.L.; Nelson Jr., O.E.

Plant Cell 3, 1337-1348, 1991

A:Title: Analysis of maize brittle-1 alleles and a defective suppressor-mutator-induced

A:Reference number: JQ1459; MUID:93005685

A:Accession: JQ1459

A:Molecule type: mRNA

A:Residues: 1-436 <SUL>

A:Cross-references: GB:M79333; NID:gl68425; PIDN:AAA33438.1; PID:gl68426

C:Comment: This protein acts as an adenylate translocator in amyloplasts.

C:Genetics:

A:Introns: 206/3; 260/3

C:Superfamily: Bt1 protein; ADP,ATP carrier protein repeat homology

C:Keywords: chloroplast; duplication; transmembrane protein

F:1-75/Domain: transit peptide (amyloplast) #status predicted <TNP>

F:76-436/Product: Bt1 protein #status predicted <MAT>

F:131-217/Domain: ADP,ATP carrier protein repeat homology <ACP1>

F:226-312/Domain: ADP,ATP carrier protein repeat homology <ACP2>

F:229-247/Domain: transmembrane #status predicted <TM1>

F:323-413/Domain: ADP,ATP carrier protein repeat homology <ACP3>

F:327-347/Domain: transmembrane #status predicted <TM2>

Query Match

Best Local Similarity 13.7%; Score 266.5; DB 2; Length 436;

Matches 98; Conservative 57; Mismatches 138; Indels 85; Gaps 16;

```
QY 19 PG-RSPGESALLDGLQRGVGGAGGEGAGACR-----PPVRQDPDSDPDYEALPAGATV 72
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 90 PGSRPPGRR-----GRGSEEEAEGRHHEAAAGSEPEEGQGQDRQAPAPA-- 136
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 73 TTHMVAGAVAGILEHCVMYPIDCVKTRMO--SLQPDPAARYRNVLREALWRIIRTEGLWRP 130
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 --RLVSGAIGAGAVSRFTVPLETIRTHLMVGSIGVDSMA---GVFQ--W-IMQNEGWTGL 188
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 131 MRGLNVTATGAGPAHALYFACYEKLKLTSDVIHPCGNH-----IANGAAGCVAT 181
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 189 FRGNVNVLRVAPSIAIEHTYDTAKKFLT----PKGDEPPPIPIPTPLVAGALAGFAST 244
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 LLHDAAMNPAEVVYKORMQNTSPYHRVTDVRAVWQNEGAGAFYRSYTTQLTMTNVPQAI 241
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 L-----CTYPMELIKRTVTIEKDVYDNVAHAFVKILDEGPSELYRGLTSLIGVVPYAC 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 242 HFMTYEFLOEHF-----NPPORRYNSSHVLVSGACAGAVAAATTPLDVCKTLTNTQESL 295
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 NEYAVETLRLYRRATGRPGADVGVPVATLLIGSRAAGATASSATFFLEVARKQM----- 355
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 296 ALNSHTGHTGHTGMASAFRTVYQ-----VGGWTAYFRGVQARVYIQIPSTAI 341
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 356 -----VGAVGG-----ROYVQNVLHAIYCIILKKEGAGGL---YRGLGPPSCIKLMPAAGI 401
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 342 ANSVVEFFKYLITKROEE 359
```

```
Db 402 AFWCYEACKKILYDKEDE 419
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: August 27, 2002, 03:57:40
Job time: 4946 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:56:08 ; Search time 37.5 Seconds
(without alignments)

375.838 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGGAGGPAAGPG.....VYEFKYLIKROEWRAGK 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	558.5	28.8	304	1 MRS4_YEAST	P3500 saccharomyc
2	555.5	28.6	314	1 MRS3_YEAST	P10566 saccharomyc
3	313.5	16.2	368	1 YN39_YEAST	Q03829 saccharomyc
4	295	15.2	328	1 Y051_CAEEL	Q09461 caenorhabdi
5	289	14.9	359	1 CG69_HUMAN	Q9bz14 homo sapien
6	272.5	14.0	307	1 ODC2_YEAST	Q92927 saccharomyc
7	268	13.8	676	1 CMC2_MOUSE	Q9qxx4 mus musculu
8	266.5	13.7	436	1 PTL_WAIZE	P29518 zea mays (m
9	265.5	13.7	284	1 PET8_YEAST	P38921 saccharomyc
10	265	13.7	675	1 CMC2_HUMAN	Q9ujs0 homo sapien
11	261.5	13.5	695	1 CMCL1_DROME	Q9va73 drosophila
12	259	13.4	320	1 DNC_HUMAN	Q9hc21 homo sapien
13	250.5	12.9	322	1 SFCL1_YEAST	P33303 saccharomyc
14	244.5	12.6	315	1 MFT_HUMAN	Q9h2d1 homo sapien
15	240	12.4	315	1 SA18_HUMAN	Q9h1k4 homo sapien
16	240	12.4	678	1 CMCL1_HUMAN	Q75746 homo sapien
17	235	12.1	312	1 SA18_MOUSE	Q9db41 mus musculu
18	230.5	11.9	345	1 YDE9_SCHPO	Q10442 schizosacch
19	229.5	11.8	311	1 TXTP_HUMAN	P53007 homo sapien
20	228	11.8	373	1 YIA6_YEAST	P40556 saccharomyc
21	227	11.7	299	1 ODC_HUMAN	Q9bqt8 homo sapien
22	225.5	11.6	366	1 YG5F_YEAST	P53320 saccharomyc
23	225	11.6	335	1 YEAG_YEAST	P39953 saccharomyc
24	224.5	11.6	311	1 TXTP_BOVIN	P79110 bos taurus
25	224.5	11.6	588	1 CMC2_CAEEL	Q20799 caenorhabdi
26	223	11.5	301	1 MCAT_HUMAN	Q43772 homo sapien
27	222.5	11.5	302	1 YDIK_SCHPO	Q10248 schizosacch
28	221	11.4	311	1 TXTP_RAT	P32089 rattus norv
29	219	11.3	271	1 YAD8_SCHPO	Q09834 schizosacch
30	219	11.3	301	1 ORTL_MOUSE	Q9wvd5 mus musculu
31	219	11.3	312	1 TXTP_CAEEL	P34519 caenorhabdi
32	215.5	11.1	301	1 MCAT_RAT	P97521 rattus norv
33	215	11.1	702	1 CMCL1_CAEEL	Q21153 caenorhabdi

34	211.5	10.9	301	1 ORT2_HUMAN	Q9bxi2 homo sapien
35	208	10.7	587	1 CMC3_CAEEL	Q19529 caenorhabdi
36	207	10.7	300	1 YEO3_YEAST	P40035 saccharomyc
37	204	10.5	311	1 MCP_YEAST	P23641 saccharomyc
38	201.5	10.4	292	1 ORT1_YEAST	Q12375 saccharomyc
39	201.5	10.4	308	1 UCP3_MOUSE	P56501 mus musculu
40	201	10.4	307	1 PM34_HUMAN	O43808 homo sapien
41	200	10.3	307	1 PM34_MOUSE	O70579 mus musculu
42	197.5	10.2	299	1 TXTP_YEAST	P38152 saccharomyc
43	196.5	10.1	307	1 UCP1_HUMAN	P25874 homo sapien
44	195.5	10.1	377	1 RIM2_YEAST	P38127 saccharomyc
45	194.5	10.0	311	1 UCP3_CANFA	Q9n219 canis famil

ALIGNMENTS

```

RESULT 1
MRS4_YEAST
ID MRS4_YEAST STANDARD; PRT; 304 AA.
AC P23500;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Mitochondrial RNA splicing protein MSR4.
GN MRS4 OR YKR052C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M1301.
RX MEDLINE=91108815; PubMed=1703236;
RA Wiesenberger G., Link T.A., von Ahlsen U., Waldherr M., Schweyen R.J.;
RT "MRS3 and MRS4, two suppressors of mtRNA splicing defects in yeast,
RL J. Mol. Biol. 217:23-37(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Visser S., Urrestarazu L.A., Jauniaux J.-C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: MRS4 SUPPRESSES A MITOCHONDRIAL SPLICE DEFECT IN THE
CC FIRST INTRON OF THE COB GENE. IT MAY ACT AS A CARRIER, EXERTING
CC ITS SUPPRESSOR ACTIVITY VIA MODULATION OF SOLUTE CONCENTRATIONS IN
CC THE MITOCHONDRION (POSSIBLY OF CATIONS). NOT ESSENTIAL.
CC !- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC !- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,
CC TO YEAST MRS3; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER
CC PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56444; CAA39828.1; -.
CC EMBL; Z28277; CAA82130.1; -.
CC PIR; S13533; S13533.
CC SGD; S0001760; MRS4.
CC InterPro; IPR001993; Mitoch_carrier.
CC Pfam; PF00153; mito_carr; 3.
CC PROSITE; PS00215; MITOCH_CARRIER; 2.
CC KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
CC mRNA processing.
CC REPEAT 9 110 BLOCK I (APPROXIMATE).
CC REPEAT 111 202 BLOCK II (APPROXIMATE).
CC REPEAT 203 304 BLOCK III (APPROXIMATE).
CC REPEAT 23 41 POTENTIAL.

```

FT TRANSMEM 83 102 POTENTIAL.
 FT TRANSMEM 120 139 POTENTIAL.
 FT TRANSMEM 175 194 POTENTIAL.
 FT TRANSMEM 209 228 POTENTIAL.
 FT TRANSMEM 275 288 POTENTIAL.
 SQ SEQUENCE 304 AA; 33307 MW; 5ABBF985B547EDC CRC64;

Query Match 28.8%; Score 558.5; DB 1; Length 304;
 Best Local Similarity 40.1%; Pred. No. 3.7e-37;
 Matches 118; Conservative 53; Mismatches 120; Indels 3; Gaps 2;

QY 62 DYALPAGATVTHMVAGAVAGILEHCVMPIDCVKTRMOSLPDPAARVNVLEALMRI 121
 ||||| : : : : ||||| : ||||| : ||||| : : : : :
 Db 13 DYALPSHAPLHESOLLAGAPAGIMEHSLMFPIDALKTRVQAAGLNKAAS-TGMISQISKI 71
 ||||| : : : : ||||| : ||||| : ||||| : : : : :
 QY 122 IRTEGLWRMRGLNVNTATGAGPAHALYFACYEKLKTLSDVIHPGNSHTANGAAGCVAT 181
 || : : ||||| : || : : : :
 Db 72 STMEGSMALWKGVQSVILGAGPAHAYVFGTYEFCARLISPEDMQTHQPMKLTALSGTIAT 131
 || : : ||||| : || : : : :
 QY 182 LLHDAAMNPAEVVKORMQMYNSPHYRVTDCVRAVWONEGAGAFYRSYTTOLTWNVPFOAI 241
 || : : ||||| : || : : : :
 Db 132 IADALMNPDTVKRLQLDTN--LRVNWTKQIYONEGPAFYYSPTTLAMNIPFAAF 189
 || : : ||||| : || : : : :
 QY 242 HFMTYEFLOHFNPPQRYNPSSHVLSGACAGAVAAATTPDYCKTLLNTQESLALNSHI 301
 : || : : ||||| : || : : : :
 Db 190 NFMIYESAKFFPNOSYNPLIHCLGGISGATCAALTPDLCIKTVLQVRGSETVSIEI 249
 : || : : ||||| : || : : : :
 QY 302 TGHITGMASAFRIVYQVGGVTAYFRGVQARVIYQIPSTAIANSVYEFFKYLTK 355
 : || : : ||||| : || : : : :
 Db 250 MKDANTFGRASRAILEVHGKGFWRGLKPRIVANIPATAISWTAYECAKFLMK 303
 : || : : ||||| : || : : : :

RESULT 2

MRS3_YEAST
 ID MRS3_YEAST STANDARD; PRT; 314 AA.
 AC P10566;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Mitochondrial RNA splicing protein MRS3.
 GN MRS3 OR YJL133W OR J0675.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IC8/R101;
 RX MEDLINE=88121698; PubMed=2448588;
 RA Schmidt C., Soellner T., Schweyen R.J.;
 RT "Nuclear suppression of a mitochondrial RNA splice defect: nucleotide
 RT sequence and disruption of a mitochondrial carrier family.";
 RL Mol. Gen. Genet. 210:145-152(1987).
 RN [2]
 RP REVISIONS, SEQUENCE FROM N.A.
 RC STRAIN=M1301;
 RX MEDLINE=91108815; PubMed=1703236;
 RA Wiesenberger G., Link T.A., von Ahsen U., Waldherr M., Schweyen R.J.;
 RT "MRS3 and MRS4, two suppressors of mtRNA splicing defects in yeast,
 RT are new members of the mitochondrial carrier family.";
 RL J. Mol. Biol. 217:23-37(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=96408771; PubMed=8813765;
 RA Katsoulou C., Tzermia M., Tavernarakis N., Alexandraki D.;
 RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast
 RT chromosome X reveals 14 known genes and 13 new open reading frames
 RT including homologues of genes clustered on the right arm of
 RT chromosome XI.";
 RL Yeast 12:787-797(1996).
 CC -!- FUNCTION: MRS3 SUPPRESSES A MITOCHONDRIAL SPLICE DEFECT IN THE

CC FIRST INTRON OF THE COB GENE. IT MAY ACT AS A CARRIER, EXERTING
 CC ITS SUPPRESSOR ACTIVITY VIA MODULATION OF SOLUTE CONCENTRATIONS IN
 CC THE MITOCHONDRION (POSSIBLY OF CATIONS).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,
 CC TO YEAST MRS4; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER
 CC PROTEIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X56445; CAA39829.1; -
 CC EMBL; X56445; CAA39830.1; ALT_INIT.
 CC EMBL; X06239; CAA29582.1; ALT_SEQ.
 CC EMBL; X87371; CAA60822.1; -
 CC EMBL; Z49408; CAA89428.1; -
 CC FIR; S01267; S01267.
 CC FIR; S20228; S20228.
 CC SGD; S0003669; MRS3.
 CC InterPro; IPR001993; Mitoch_carrier.
 CC Pfam; PF00153; mito_carr; 3
 CC PROSITE; PS00215; MITOCH_CARRIER; 2.
 CC Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 CC mRNA processing.
 FT REPEAT 19 120 BLOCK I (APPROXIMATE).
 FT REPEAT 121 212 BLOCK II (APPROXIMATE).
 FT REPEAT 213 314 BLOCK III (APPROXIMATE).
 FT TRANSMEM 33 52 POTENTIAL.
 FT TRANSMEM 93 112 POTENTIAL.
 FT TRANSMEM 130 149 POTENTIAL.
 FT TRANSMEM 185 204 POTENTIAL.
 FT TRANSMEM 219 238 POTENTIAL.
 FT TRANSMEM 285 298 POTENTIAL.
 SQ SEQUENCE 314 AA; 34501 MW; 861145CE6AEF321 CRC64;

Query Match 28.6%; Score 555.5; DB 1; Length 314;
 Best Local Similarity 39.1%; Pred. No. 6.6e-37;
 Matches 118; Conservative 50; Mismatches 131; Indels 3; Gaps 2;

QY 52 PVQRDPDGPDYALPAGATVTHMVAGAVAGILEHCVMPIDCVKTRMOSLPDPAARY 111
 : : ||||| : : ||||| : ||||| : ||||| : ||||| :
 Db 13 PIPAIPMDLPDYALPHTAPLYHQLIAGAPAGIMEHSMPEPIDALKTRIOSANAKLSA- 71
 : : ||||| : : ||||| : ||||| : ||||| : ||||| :
 QY 112 RNVLEALWRIIRTEGLWRPMRGLNVNTATGAGPAHALYFACYEKLKTLSDVIHPGNSHI 171
 : || : : || : : || : : ||||| : || : : : :
 Db 72 KNMLSQISHISTEGTLALWKGVQSVILGAGPAHAYVFGTYEFCCKNLIDSSDTQTHHPF 131
 : || : : || : : || : : ||||| : || : : : :
 QY 172 ANGAAGCVATLLHDAAMNPAEVVKORMQMYNSPHYRVTDCVRAVWONEGAGAFYRSYTTQ 231
 : || : : ||||| : || : : : :
 Db 132 KYAISGACATTASDALMNPDTIKRIQLNTSA--SVWTKQIYQSEGLAARFYSPPT 189
 : || : : ||||| : || : : : :
 QY 232 LFMNVFQAIHMTYEFLOHFNPPQRYNPSSHVLSGACAGAVAAATTPDYCKTLLNT 291
 : ||||| : : : : ||||| : ||||| : ||||| : ||||| : ||||| :
 Db 190 LVNIPFAAFNFVYESSTKFLNPSNEYNPLIHCLGSGISGSCAAITTPDLCIKVLQI 249
 : || : : ||||| : || : : : :
 QY 292 QESLALNSHITGHTGMASAFRTVYQVGGVTAYFRGVQARVIYQIPSTAIANSVYEFFKY 351
 : || : : ||||| : || : : : :
 Db 250 RGSQTVSLEIMRKADTFSKAASAIQYGVKGFWRGKPRIVANMPATISWTAYEACAKH 309
 : || : : ||||| : || : : : :
 QY 352 LI 353
 :
 Db 310 FL 311

RESULT 3
 YM39_YEAST

```
ID YM39_YEAST STANDARD; PRT; 368 AA.
AC Q03829;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative mitochondrial carrier YMR166C.
GN YMR166C OR YMR520.15C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (Potential).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 249705; CAA89802.1; -
DR SGB; S0004776; YMR166C.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
SQ SEQUENCE 368 AA; 40992 MW; B583100018DF045D CRC64;
```

Query Match 16.2%; Score 313.5; DB 1; Length 368;
Best Local Similarity 29.0%; Pred. No. 1.1e-17;
Matches 88; Conservative 66; Mismatches 114; Indels 35; Gaps 11;

```
QY 76 MVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYNVLEALWRIIRTEGLWRPMLN 135
DB 57 VVSGIGGKIGDSAMHSLDTVKTRQQG-APN-VKRYRNMSAYRTIWLBEGVRRGLYGY 114
QY 136 VTA-TGAGPAHALYFACYEKLKLTSDVIHPGNGSHIANGAAGCVATLHDAAMNPAEV 194
DB 115 MAALGSPSPAIFGTGYEYKRTMEDWQI--NTITHLSAGFLGDFISSFVYPSEVL 172
QY 195 KORMQM-----YHRVTDVRAVQWNGAGAFYSYTTQLTMNVFQAIHFM 244
DB 173 KTRLOQGRFNPFQSGYNYNLRNAIKTVKEGFRSLFFGYKATLARDLPFSALQFA 232
QY 245 TYE-FLOEHFNQRR-----YNPSSHVLSGACAGAVAAAATPLDVKCTLLNTQESLA 297
DB 233 FYEERQALAFKEQDGRDGLSIPNEILTGACAGLAGIITPMDVVKVTRTQOQPSQ 292
QY 298 NS-----HIT-GHITGMAS-----AFRTVYQGVGTAFYFRGVOARVYIQIPSTAIWS 344
DB 293 SNKSYSVTHPHVTNGRPAALNSISLRTVYQSEGLGFFSGVGPRFVMTVSQSIMLL 352
QY 345 VYE 347
DB 353 LYQ 355
```

RESULT 4
YQ51_CAEEL

```
ID YQ51_CAEEL STANDARD; PRT; 328 AA.
AC Q09461;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative mitochondrial carrier C16C10.1.
GN C16C10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Lloyd C.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 246787; CAA86739.1; -
DR WormPep; C16C10.1; CE01489.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein; Mitochondrion; Repeat; Transmembrane;
KW Transport.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
SQ SEQUENCE 328 AA; 36743 MW; EA1E9E329A764DF6 CRC64;
```

Query Match 15.2%; Score 295; DB 1; Length 328;
Best Local Similarity 27.8%; Pred. No. 2.8e-16;
Matches 101; Conservative 48; Mismatches 146; Indels 68; Gaps 14;

```
QY 22 SPGESALLDGLQRGVGRGAGGEGAGACRPP-----VRQDPDSGPDYEALPAGATVTHM 76
DB 3 SPES-----GKTTNCSGAPSSSCVTPLDVVKIRLQQQTRP----FPGKECFYH- 48
QY 77 VAGAVAGILEHCVMYPIDCVKTRMQSLQ-----DPAARYNVLEALWRIIRTEGLWRPMLR 132
DB 49 -----NGLMEHY-----CVSCEVR--KPEWYQRPNGFRGTADAIVKIHREGIRSLWS 95
QY 133 GLNVTATGAGPAHALYFACYEKLKLTSDVI-----HPGNGSHIANGAAGCVAT 181
DB 96 GLSPFMVMPALPATVFTYTDNLVSWLKKMKCRRRAFSPEKTPPDWS--AAVAGIAR 153
QY 182 LLHDAAMNPAEVYKORMQMINSPYHRVTDVRAVQWNGAGAFYSYTTQLTMNVFQAI 241
DB 154 TIATVTVSPFIEMIRTKMQSKRLTYHEIGHVRSRMATKGISSPYLGTPTMLRDIFSGI 213
QY 242 HMTVEFFLOEHFNQRRYNPSSH-----VLSGACAGAVAAAATPLDVKCTLLNTQESLA 296
DB 214 YWAGYDLFT--NLQRRQGDHPNFVVSFVSGAAAGVWASIFTHPPDVIKT--NCO---- 265
QY 297 LNSHTTGHTGTGNASAFRTV---YQVGGVTAFYFRGVOARVYIQIPSTAIWSYEFKYL 352
DB 266 --IRGGSIDDMNKSIITTVIKDMYHSRGISAFSSGLVPRLVKVPSPSCAIMSIFYEFKFL 323
QY 353 ITK 355
DB 324 FQK 326
```

RESULT 5

CG69_HUMAN
ID CG69_HUMAN STANDARD; PRT; 359 AA.
AC Q9B274; O9Y379; O9P182;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial carrier protein CGI-69 (PRO2163).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SUBCELLULAR LOCATION, AND
RP TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX PubMed=11139402;
RA Yu X.X., Lewin D.A., Zhong A., Brush J., Schow P.W., Sherwood S.W.,
RA Pan G., Adams S.H.;
RT "Overexpression of the human 2-oxoglutarate carrier lowers
RT mitochondrial membrane potential in HEK-293 cells: contrast with the
RT unique cold-induced mitochondrial carrier CGI-69.";
RL Biochem. J. 353:369-375(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Dusterhoef A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isosai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain, and Colon;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 15-359 FROM N.A. (ISOFORM 1).
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
RA He F.;
RT "Functional prediction of the coding sequences of 79 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1/CGI-69L (shown here) and 2;
CC are produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: Expressed in many tissues. Abundant in testis
CC and kidney.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF317711; RAG60687.1; -
DR EMBL; AF151827; AAD34064.1; -
DR EMBL; AL133584; CAB63728.1; -
DR EMBL; AK026060; BAB15341.1; -
DR EMBL; BC001398; AAH01398.1; -
DR EMBL; BC009330; AAH09330.1; -
DR EMBL; AF119864; RAG69618.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
KW Alternative splicing; Polymorphism.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT VARSPLIC 64 72 LPSSLOSTG -> W (IN ISOFORM 2).
FT VARIANT 247 247 L -> F.
FT
FT CONFLICT 266 266 /FTID=VAR_012756.
FT SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;
SQ

Query Match 14.9%; Score 289; DB 1; Length 359;
Best Local Similarity 24.5%; Pred. No. 9.4e-16;
Matches 89; Conservative 50; Mismatches 140; Indels 84; Gaps 8;

QY 67 PAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQDPDA----- 109
DB 6 PAGISPLQOMVAGSGTGAVTSLFMTPLDVVKVRLQSORPSMASLAPSSRLWSLSYTKLP 65
QY 110 -----RYRNVEALWRIIRTEG---LWR 129
DB 66 SSLSQSTGKCLLYCNGVLEPLLYLCPNGRCATWFDPTFRFTGMDAFVKYIHRHGTRLW- 124
QY 130 PMRGLNVATGAGPAHALYFACYEKLKLT-----SDVIHPGGNSHIANGAACGVATLL 183
DB 125 --SGLPATLVMTVPATAIYFTAYDQLKAFLCGRALTSGLYAP-----MWAGALARLG 174
QY 184 HDAAMNPAEVVKORMOMYNSPYHRVTDCVRVWQNEGAGAFYSYTTOLTMTNVPFQAIHF 243
DB 175 TVTVISPLELMRTKLOAQHVSYRELGCACVRTAVAQGGWRSRLWLGWGTALRDVPFSALYW 234
QY 244 MTVEFLQEHFN---PQRRYNPSSHVLSGACAGAVAAATPLDVCKT-----LLNTOESLA 296
DB 235 FNYELVKSWLNGLRPKDQTSVGMFSVAGGISGVIAVAVLTLPFDVVKTORVALGAEAVR 294
QY 297 LNSHITGHITGMASAFRTVYQGVGTAYFRGVOARVIYQIPSTAIAMSVYEFFKYLITKR 356
DB 295 VNPL---HVDSTWLLLRIRAESGTGLFAGFLPRIIKAAPSCAIMISTVEFGKSPFQRL 351
QY 357 QEE 359
DB 352 NQD 354

RESULT 6
ODC2_YEAST ID ODC2_YEAST STANDARD; PRT; 307 AA.
AC Q99297;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial 2-oxodicarboxylate carrier 2.

GN ODC2 OR YOR222W OR YOR50-12.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RA Galissson F., Dujon B.;
 RT "Sequence and analysis of a 33 kb fragment from the right arm of
 RL chromosome XV of the yeast Saccharomyces cerevisiae.";
 RL Yeast 12:877-885(1996).
 RN [2]
 RP CHARACTERIZATION.
 RX PubMed=11013234;
 RA Palmieri L., Agrimi G., Runswick M.J., Fearnley I.M., Palmieri F.,
 RA Walker J.E.;
 RT "Identification in Saccharomyces cerevisiae of two isoforms of a novel
 RL mitochondrial transporter for 2-oxoadipate and 2-oxoglutarate.";
 RL J. Biol. Chem. 276:1916-1922(2001).
 CC -!- FUNCTION: Transports C5-C7 oxodicarboxylates across the inner
 CC membranes of mitochondria. Can transport 2-oxoadipate, 2-
 CC oxoglutarate, adipate, glutarate, 2-oxopimelate, oxaloacetate,
 CC citrate and malate. The main physiological role is probably to
 CC supply 2-oxoadipate and 2-oxoglutarate from the mitochondrial
 CC matrix to the cytosol where they are used in the biosynthesis of
 CC lysine and glutamate, respectively, and in lysine catabolism.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; 275130; CAA99440.1; -;
 DR EMBL; X92441; CAA63185.1; -;
 DR SGD; S0005748; ODC2.
 DR InterPro; IPR002067; Mit_carrier.
 DR InterPro; IPR001993; Mitoch_carrier.
 DR Pfam; PF00153; mito_carr; 3.
 DR PRINTS; PR00926; MITOCH_CARRIER.
 DR PROSITE; PS00215; MITOCH_CARRIER; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 280 300 POTENTIAL.
 SQ SEQUENCE 307 AA; 34007 MW; 4089082A64DBA97C CRC64;

Query Match 14.0%; Score 272.5; DB 1; Length 307;
 Best Local Similarity 27.2%; Pred. No. 1.6e-14;
 Matches 81; Conservative 54; Mismatches 130; Indels 33; Gaps 7;
 QY 77 VAGAVAGILEHCVMYPIDCVKTRMOSLPDPA-----RYRNVLALWRIIRTEGLW 128
 DB 17 ISGAVAGISELVMTPLDVKVTRFQLEVTPTAAVGVKQVERINGVDCLKIKVKEGFS 76
 QY 129 RPRMGLNVTATGAGFAHALYFACYEKLKTKLSDVHPGNSHANGAACVATLLHDAAM 188
 DB 77 RLYRGISSPMLMEAPKATKFCACNDYOYKIFKNLENTNETTKISIAAGASAGMTAAVI 136
 QY 189 NPAEVVVKQRMQNSGYHVRVTCVRAVWONEGAGAFYRYYTTLQTMNVFPQAIHF-MTVE 247
 DB 137 VPFEIKIRMQDVKSSYGLPMDCLKTKTKNEGIMGLYKGIESTMNRNALWNGGYFCVIYQ 196

QY 248 FLOEHFNPPORRYNP-----SSHLVSGACAGAVAAAAATTPLDVCKTLLNTQESIALN 298
 DB 197 V-----RNSMPVAKTKQKTRNDLIAGAIIGTGTMLNTFFDVVKSHIQSDV--AVS 246
 QY 299 SHITGHITGMASAFRTVYQVGVYATYFRGVOARVYVOIPSTAIWSVY-----EFFKYL 352
 DB 247 SAVKKNWCLPSLL-VIYREGFRALYKGFVPKVCRLAPGSGSLMLVVFVTGMNFFRDL 303
 RESULT 7
 CMC2_MOUSE STANDARD; PRT; 676 AA.
 ID CMC2_MOUSE Q9QXX4; Q9DCF5; Q9CF6;
 AC Q9QXX4; Q9DCF5; Q9CF6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Calcium-binding mitochondrial carrier protein Aralar2 (Solute carrier
 DE family 25, member 13) (Citrin).
 GN SLC25A13 OR ARALAR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=20079163; PubMed=10610724;
 RA Sinasac D.S., Crackower M.A., Lee J.R., Kobayashi K., Sahaki T.,
 RA Scherer S.W., Tsui L.-C.;
 RT "Genomic structure of the adult-onset type II citrullinemia gene,
 RT SLC25A13, and cloning and expression of its mouse homologue.";
 RL Genomics 62:289-292(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo, and Kidney;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein S.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momabaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Willing L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -!- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER. MAY HAVE
 CC A FUNCTION IN THE UREA CYCLE (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: AT E10.45, EXPRESSED IN BRANCHIAL ARCHES AND
 CC LRTMB AND TAIL BUDS. AT E13.5 EXPRESSION IS PREDOMINANT IN
 CC EPITHELIAL STRUCTURES AND THE FOREBRAIN, KIDNEY AND LIVER.
 CC EXPRESSION IN LIVER IS MAINTAINED INTO ADULTHOOD.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR
 CC SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF164632; AAF21426.1; -.
CC EMBL; AK002829; BAB22390.1; -.
CC EMBL; AK012670; BAB28397.1; -.
CC MGD; MGI:1354721; SIC25a13.
CC DR InterPro; IPR002048; EF-hand.
CC DR InterPro; IPR002067; Mit_carrier.
CC DR InterPro; IPR001993; Mitoch_carrier.
CC DR Pfam; PF00036; ehand; 3.
CC DR Pfam; PF00153; mito_carr; 3.
CC DR PRINTS; PR00926; MITOCARRIER.
CC DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
CC DR PROSITE; PS00215; MITOCH_CARRIER; 3.
CC KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
CC KW Calcium-binding.
CC FT TRANSMEM 333 350 1 (POTENTIAL).
CC FT TRANSMEM 394 413 2 (POTENTIAL).
CC FT TRANSMEM 437 450 3 (POTENTIAL).
CC FT TRANSMEM 486 505 4 (POTENTIAL).
CC FT TRANSMEM 525 542 5 (POTENTIAL).
CC FT TRANSMEM 582 601 6 (POTENTIAL).
CC FT CA_BIND 66 77 EF-HAND 1.
CC FT CA_BIND 100 111 EF-HAND 2.
CC FT DOMAIN 139 150 ANCESTRAL CALCIUM SITE 3.
CC FT CA_BIND 171 182 EF-HAND 4.
CC FT REPEAT 207 422 1.
CC FT REPEAT 423 512 2.
CC FT REPEAT 513 615 3.
CC FT CONFLICT 312 312 MISSING (IN REF. 2; BAB28397).
CC FT CONFLICT 442 442 S -> F (IN REF. 2; BAB22390).
CC SQ SEQUENCE 676 AA; 74466 MW; 5E3952F08E5E58C5 CRC64;

Query Match 13.88; Score 268; DB 1; Length 676;
Best Local Similarity 28.24; Pred. No. 9.2e-14;
Matches 82; Conservative 47; Mismatches 122; Indels 40; Gaps 9;

QY 79 GAVAGILEHCVMYPIDCVKTRMQSLQDPDA-----ARYRNVLALWRIIRTEGLWRPGRGL 134
DB 336 GSIAGAGATAYPIDLVKTRMQNRQSGFVGMELMKNSEDFCKKVLVRGEGFGLYRGL 395
QY 135 NVATGAGPAHALYFACYEKLKTLSDVI-----HPGNSHI-----ANGAAGCVATLLH 184
DB 396 LPQLLGVAPEKA-----IKLTNDFVRDKFMKDGSPVLLAEIFAGGCGAGSVIF- 446
QY 185 DAANPAEVVKORMQNSPYHRVTDVCR-----AVWNEGAGAFYRSYTTQLTMNVFQA 240
DB 447 ---TNPLEIVKIRLQVAG-----EITTPRVSAVSVRDLGFFGIYKGAACFLRDIPFSA 499
QY 241 IHFWTYEFLQEHF--NPQRRYNPSSHVLGACAGAVAAATTPLDVCKTLTLLNTOESLNS 299
DB 500 IYFCYARHVKASFNAGQVSPGSLLAGAAGNPAASLVTTPADVIKTRLOVARAGQTT 559
QY 300 HITGHITGMSAFRTYQVGGVTAIFRGVQARVYIQTIPSTAIWSVYEFK 350
DB 560 Y-----NGVTDCTFRKILREEGPRALMKVGAARFSSPQGVQVTLTLLYELLQ 605

RESULT 8
BT1_MAIZE STANDARD; PRT; 436 AA.
AC P29518;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Brittle-1 protein, chloroplast precursor.
GN BT1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005685; PubMed=1668652;
RA Sullivan T.D., Strelow L.I., Illingworth C.A., Phillips R.L.,
RA Nelson O.E. Jr.;
RT *Analysis of maize brittle-1 alleles and a defective
RT Suppressor-mutator-induced mutable allele.*;
RL Plant Cell 3:1337-1348(1991).
CC -|- FUNCTION: COULD PLAY A ROLE IN AMYLOPLAST MEMBRANE TRANSPORT.
CC -|- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
CC -|- TISSUE SPECIFICITY: ENDOSPERM OF DEVELOPING KERNELS.
CC -|- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M79333; AAA33438.1; -.
CC PIR; JQ1459; JQ1459.
CC MaizeDB; 47578; -.
CC InterPro; IPR002067; Mit_carrier.
CC InterPro; IPR001993; Mitoch_carrier.
CC Pfam; PF00153; mito_carr; 3.
CC PRINTS; PR00926; MITOCARRIER.
CC PROSITE; PS00215; MITOCH_CARRIER; 1.
CC TRANSIT peptide; Chloroplast; Amyloplast; Transmembrane.
CC TRANSIT 1 75 CHLOROPLAST (POTENTIAL).
CC CHAIN 76 436 BRITTLE-1 PROTEIN.
CC TRANSMEM 229 247 POTENTIAL.
CC TRANSMEM 327 347 POTENTIAL.
CC SQ SEQUENCE 436 AA; 46627 MW; 960C05F603E9DAE CRC64;

Query Match 13.7%; Score 266.5; DB 1; Length 436;
Best Local Similarity 25.9%; Pred. No. 7.1e-14;
Matches 98; Conservative 57; Mismatches 138; Indels 85; Gaps 16;

QY 19 PG-RSPGESALLDGLQVRGVRGAGGGEAGACR-----PPVRQDPDSDGPDYEALPAGATV 72
DB 90 PGSRPPGRR-----GRGSEEEEAERHHEEAAAAGRSPEEGQGDQRPAPA-- 136
QY 73 TTHMVACAVAGILEHCVMYPIDCVKTRMQ--SLOPDPAARYRNVLALWRIIRTEGLWRP 130
DB 137 --RLVSGAIAAGAVSRTPVAPLETIRTHLMVSGIGVDSMA---GVFQ--W-IMQEGWTGL 188
QY 131 MRGLNVATGAGPAHALYFACYEKLKTLSDVIHPGNSH-----IANGAAGCVAT 181
DB 189 FRGNVNVLRVAPSKAIEHTYDTAKKFLT---PKGDEPKPIPTPLVAGALAGFAST 244
QY 182 LLHDAANPAEVVKORMQNSPYHRVTDVCRVAVWNEGAGAFYRSYTTQLTMNVFQA 241
DB 245 L-----CTYPMELIKTRVTIEKVDNVAHAFVKILRDEGPESELYGLTSLIGVVPYAA 300
QY 242 HEMTYEFLQEHF-----NPQRRYNPSSHVLGACAGAVAAATTPLDVCKTLTLLNTOESL 295
DB 301 NFYATETLKRILYRRATGRRPGADVGVATLLIGSAGAIASSATTFPLEVARKQM 355
QY 296 ALNSHITGHITGMSAFRTYQ-----YGGVTAYFRGVQARVYIQTIPSTAI 341
DB 356 -----VGAVGG-----RQVYQNVLHAIYILKKEGAGGL---YRGLGPSICKLMPAAGI 401
QY 342 AWSVYEFKILYTRQEE 359
DB 402 AFMCYEACKKILVDKEDE 419

RESULT 9
PET8_YEAST STANDARD; PRT; 284 AA.
ID PET8_YEAST
```



```
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF118838; AAD38501.1; -
DR EMBL; Y17571; CAB62206.1; -
DR EMBL; AF164530; AAF28473.1; -
DR EMBL; AF164525; AAF28473.1; JOINED.
DR EMBL; AF164526; AAF28473.1; JOINED.
DR EMBL; AF164527; AAF28473.1; JOINED.
DR EMBL; AF164528; AAF28473.1; JOINED.
DR EMBL; AF164529; AAF28473.1; JOINED.
DR EMBL; AC002450; AAB67049.1; -
DR EMBL; AC002540; AAB70112.1; -
DR MIM; 603859; -
DR MIM; 603471; -
DR MIM; 605814; -
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR002067; Mit_carrier.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00036; efhand; 3.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; P00926; MITOCARRIER.
DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
KW Calcium-binding; Polymorphism.
FT TRANSEM 332 349 1 (POTENTIAL).
FT TRANSEM 393 412 2 (POTENTIAL).
FT TRANSEM 436 449 3 (POTENTIAL).
FT TRANSEM 485 504 4 (POTENTIAL).
FT TRANSEM 524 541 5 (POTENTIAL).
FT TRANSEM 581 600 6 (POTENTIAL).
FT CA_BIND 66 77 EF-HAND 1.
FT CA_BIND 100 111 EF-HAND 2.
FT DOMAIN 139 150 ANCESTRAL CALCIUM SITE 3.
FT CA_BIND 171 182 EF-HAND 4.
FT REPEAT 207 421 1.
FT REPEAT 422 511 2.
FT REPEAT 512 614 3.
FT CONFLICT 231 232 EL -> VH (IN REF. 2).
FT CONFLICT 311 311 Q -> QQ (IN REF. 4).
FT CONFLICT 532 532 M -> T (IN REF. 2).
SQ SEQUENCE 675 AA; 74175 MW; AD07EDBC6C68989B CRC64;

Query Match 13.7%; Score 265; DB 1; Length 675;
Best Local Similarity 28.2%; Pred. No. 1.6e-13;
Matches 82; Conservative 47; Mismatches 122; Indels 40; Gaps 9;

QY 79 GAVAGILEHCVMYPCDKTRMSIQDPDA-----ARYRNVLALWRIETGLWRPVRGL 134
DB 335 GSVAGAGATVATYDILVTKRMQNRGTGTFVGVBELTKNSFDCPKVLRVYEGFGLYRGL 394
QY 135 NVATGAGPAHALYFACYEKLKTLSDVI-----HPGNGN-----SHIANGAAGCVATLLH 184
DB 395 LPQLLGAPEKA-----IKLTNVDVDRKFMHKDGSVPLAEILAGGAGGSQVIF- 445
QY 185 DAAMPNPAEVKQRMQNSPYHRVTDVCR-----AVWQNEGAGAPRYTTLTNVVPFOA 240
DB 446 ---TNPLEIVKIRLQVAG-----EITTPRVVSALSVVRDLGFGFIYKAKACFLRDIPFSA 498
QY 241 IHEWTFEFLQEHF-NPQRRYNPSSHVLSGACAGAVAAATTPDLVCKTLLNTQESIALNS 299
DB 499 IYPCYAHVRFASFANEDQGVSPGSLLAGAIAAGMPAASLTPADVIKTRLQVAARAGQTT 558
```

```
QY 300 HITGHITGMASAFRTVYQVGVTAYFRGQVQARVIYQIPSTAIASVYEFFK 350
DB 559 Y-----SGVIDCFKILREEPKALWKAGARVFRSSPQFGVTLITVELLQ 604

RESULT 11
CMCL_DROME
ID CMCL_DROME STANDARD: PRT; 695 AA.
AC Q9VA73; Q9VA72; Q9VA74; Q9U5V8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-binding mitochondrial carrier Aralar1.
GN ARALAR1 OR CG2139.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE-Ovary;
RX MEDLINE=20115110; PubMed=10642534;
RA Del Arco A., Agudo M., Satrustegui J.;
RT "Characterization of a second member of the subfamily of calcium-
binding mitochondrial carriers expressed in human non-excitable
tissues.";
RL Biochem. J. 345:725-732(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durlin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2195-2195(2000).
CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER (BY
SIMILARITY).
```



```

DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR002067; Mit_carrier.
DR Pfam: PF00153; mito_carr. 3.
DR PRINTS: PR00926; MITOCARRIER.
DR PROSITE: PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
SQ SEQUENCE 320 AA; 35511 MW; 57CE0F01D538B1BE CRC64;

Query Match 13.4%; Score 259; DB 1; Length 320;
Best Local Similarity 24.7%; Pred. No. 1.9e-13;
Matches 78; Conservative 62; Mismatches 134; Indels 42; Gaps 12;

QY 63 YEALPAGATVTHM--VAGAVAGILEHCVMYPIDCVKTRMOSLO-----PDPAARYRN 114
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 YDKPRDGRNNTKFOVAVAGSVGLVTRALISPFQVIRFO-LQHERLSRSDPSAKYHGI 62

QY 115 LEALWRIIRTEG---LRRPMRGLNVTATGAGPAHALYFACYEKLKLTSDVIHPGG--- 167
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 LQASRQILQEGGPAFWKHGVHPAQLISIGYAVQFLSF-----EMTELVLHVRGSYDA 115

QY 168 ---NSH-IANGAGCATLLHDAAMNPAEVVVKRMQYNP--YHRVTDVCVRVWQNEGA 221
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 116 REFVSHFVCGGLAACMATL---TVHPVDVLTFRFAQGEKPVNTLRHAGVTYRSEGP 171

QY 222 GAFYRSYTTQLTNNVFOAIFHMYEFLQEHFN-----PQRYNPNSSHVLSGACAGAVAA 276
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 172 QVYFKGLAPTILAIFFPYAGLQFSCYSLKHLKYKWAIPAEGKKNNLQNLCCSGAGVISK 231

QY 277 AATPLDVCCKTLNT---QESLALNSHITGHTGMSAFRTVYQVGVYAFRGVQARVI 333
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 232 TLTPPLDLFRKLQGVGFHARAFAFGVRRY-KGLMPCAKQVLQKQEGALGFFGLSPSL 290

QY 334 YQIPSTAIANSVYEFF 349
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 291 KAALSTGFMEFSYEFF 306

RESULT 13
SFCL_YEAST STANDARD; PRT; 322 AA.
AC P33303;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Succinate/fumarate mitochondrial transporter (Regulator of acetyl-CoA
DE synthetase activity).
DE SFCL OR ACRI OR YJR095W OR J1921.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=94203187; PubMed=7908717;
RA Fernandez M., Fernandez E., Rodicio R.;
RT "ACRI, a gene encoding a protein related to mitochondrial carriers,
RT is essential for acetyl-CoA synthetase activity in Saccharomyces
RT cerevisiae";
RL Mol. Gen. Genet. 242:727-735(1994).
RN [2]
RN SEQUENCE FROM N.A.
RA Ramezani Rad M., Kirchath L., Hollenberg C.P.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RN CHARACTERIZATION.
RP MEDLINE=98055465; PubMed=9395087;
RA Palmieri L., Lasorsa F.M., De Palma A., Palmieri F., Runswick M.J.,

```

```

Walker J.E.;
"Identification of the yeast ACRI gene product as a succinate-fumarate
transporter essential for growth on ethanol or acetate.";
FEBS Lett. 417:114-118(1997).
CC -!- FUNCTION: transports cytoplasmic succinate, derived from
CC isocitrate by the action of isocitrate lyase in the cytosol, into
CC the mitochondrial matrix in exchange for fumarate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial.
CC -!- INDUCTION: BY ETHANOL OR ACETATE AS SOLE CARBON SOURCES. REPRESSED
CC BY GLUCOSE.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z25485; CAA80973.1; -.
CC EMBL: Z49595; CAA89624.1; -.
CC PIR: S36407; S36407.
CC PIR: S43280; S43280.
CC SGD: S0003856; SFCL1.
CC InterPro: IPR001993; Mitoch_carrier.
CC Pfam: PF00153; mito_carr. 3.
CC PROSITE: PS00215; MITOCH_CARRIER; 3.
KW Transmembrane; Transport; Repeat.
FT CONFLICT 300 322
   VREHLENLGIKKNDTPKPKPLK -> RKGAFOKIWIYSRR
   MTHQSQSH (IN REF. 1).
FT
SQ SEQUENCE 322 AA; 35340 MW; EC29718A0F5011A7 CRC64;

Query Match 12.9%; Score 250.5; DB 1; Length 322;
Best Local Similarity 26.3%; Pred. No. 9.1e-13;
Matches 83; Conservative 44; Mismatches 122; Indels 67; Gaps 12;

QY 75 HMVAGAVAGILEHCVMYPIDCVKTRMOSLOPDPAARYRN-----LEALWRIIR 123
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 NLMAAGTAGLFEALCCHPDITKVRMQI-----YRRVAGIEHVPPGFIKTRTIYQ 64

QY 124 TEGLRPMRGLNVTATGAGPAHALYFACYEKLKLT-----SDVIHPGNSHANGAGCV 179
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 65 KEGFLALYKGLGAVVIGIIPKMAIRFSSYEFYRTLLYNKESGIVST-GNTEVAGVGAGIT 123

QY 180 ATLLHDAAMNPAEVVVKORMQ-----MYNSPYHRVTCVRAVWQNEGAGAFYR 226
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 EAVL---VVPNMEVVKIRLQAHLTPSEPNAQPKYNNAIHAAATYIVK----EEGVSAIYR 176

QY 227 SYTTQLTMNVPFOAIHEPTY---EFLQEHFNPPORRYNPSSHV-LSGACAGAVAAATTP 281
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 177 GVSITAAQATNQANFTVYSKLEFLQNH--OMDVLPSWETSCIGLISGIPFSNAP 234

QY 282 LDVCKTLINTQESLALNSHITGHTGMSAFRTVYQVG-----GVTAYFRGVQARVY 334
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 235 LDTIKTRLQDKSISLEKQ-----SGMKIITITGAQLLKEGFRALYKYGITPRVMR 285

QY 335 QIPSTAIANSVYEFFK 350
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 286 VAPGQAVTFTVYEVVR 301

RESULT 14
MFT_HUMAN
ID MFT_HUMAN STANDARD; PRT; 315 AA.
AC Q9H2D1;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial folate transporter/carrier.
GN MFT.

```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538421; PubMed=10978331;
RA Titus S.A., Moran R.G.;
RT "Retrovirally mediated complementation of the glyB phenotype. Cloning
RT of a human gene encoding the carrier for entry of folates into
RT mitochondria."
RL J. Biol. Chem. 275:36811-36817(2000).
CC -!- FUNCTION: Transport folate across the inner membranes of
CC mitochondria.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF283645; AAG37834.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PF00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 89 106 POTENTIAL.
FT TRANSMEM 227 243 POTENTIAL.
FT TRANSMEM 281 300 POTENTIAL.
SQ SEQUENCE 315 AA; 35388 MW; 60EBCDD61951EB6E CRC64;

Query Match 12.6%; Score 244.5; DB 1; Length 315;
Best Local Similarity 24.8%; Pred. No. 2.7e-12;
Matches 79; Conservative 54; Mismatches 142; Indels 43; Gaps 11;

QY 68 AGATVTH-----MVAGAVAGILEHCVMYPCDKVTRMQ-----SLQDPDPAARYRNVL 117
DB 13 AMSTVFRHVRVYENLIAGVSGVLSNLALHPLDLVKIRFAVSDGLELRP-----KYN 68

QY 118 LWRIITRGLRMPRLNVTAGAPAHALYFACYEKLK--KTLSDVIHPGNGSHIANGA 175
DB 69 LTTIKLDGLRGLYGVTPNITWAGLSGLGFFFTNAIKSYKTEGRAHLEATEYLVSA 128

QY 176 AGCVATLLHDAAMNPAEYVVKORMQ-----YNSP---YHRVTDVCRVAVWQNEGAGAFY 227
DB 129 EAGAMTL---CITNPLWTKRLMLQYDAVNVSPHRQYKGMFDTLVKIYKYGVRGLYKG 185

QY 228 YTTQTMNVPQAIHFMTYEFLO-----EHEFN--PQRRYNPSHVLGACAGAVAAATTP 281
DB 186 FVPGI--FCTSHGALQFMAYELLKLYNOHINRLPEAQLSTVEYISVAALSKEIFAATYP 244

QY 282 LDVCKTLTNTQESLALNSHITGHTMASAFRTVYQVGVTAFFRGVQARVYQIPSTAI 341
DB 245 YQVVRARLQDQHM-----YSGVIDVTIKWRKEGVGGYFKGIAPNLIRVTPACCI 295

QY 342 AWSVVEFFKYLITRKEE 359
DB 296 TFWVYVNSHFLDLREK 313

RESULT 15
SA18_HUMAN STANDARD; PRY; 315 AA.
ID SA18_HUMAN
```

```
AC O9HLK4.
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Solute carrier family 25, member 18.
DE SLC25A18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21275466; PubMed=11381032;
RA Footz T.K., Brinkman-Mills P., Banting G.S., Maier S.A., Riazzi A.M.,
RA Bridgland L.J., Hu S., Birren B., Minoshima S., Shimizu N., Pan H.,
RA Nguyen T., Fang F., Fu Y., Ray L., Wu H., Shauli S., Phan S., Yao Z.,
RA Chen F., Huan A., Hu P., Wang Q., Loh P., Qi S., Roe B.A.,
RA McDermid H.E.;
RT "Analysis of the cat eye syndrome critical region in humans and the
RT region of conserved synten in mice: a search for candidate genes at
RT or near the human chromosome 22 pericentromere."
RL Genome Res. 11:1053-1070(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY008285; AAG22855.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 225 245 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
SQ SEQUENCE 315 AA; 33848 MW; B23A9E5036671634 CRC64;

Query Match 12.4%; Score 240; DB 1; Length 315;
Best Local Similarity 25.2%; Pred. No. 6e-12;
Matches 77; Conservative 53; Mismatches 136; Indels 40; Gaps 9;

QY 71 TTTTTHWAGVAGILEHCVMYPCDKVTRMQSLQDPDPAARYRNVLALRIITRGLWRP 130
DB 7 SITAKLVGVAGLVGVTCVFPDIDAKTRLQN--QHGRAMYKGMIDCLMKTRABGFTQM 64

QY 131 MRGLNVTATGAPAHALYFACYEKLKTLSDVIHPGNGSHI--ANGAAGCVATLLHDAAMN 189
DB 65 YRGAAVNLTLVTPERAIKLAANDFFRLL---MEDGMQRNLKWEMLAGCGAGCQVWVTC 121

QY 190 PAEVVKORMQ-----MYNSPYHRVTDVCRVAVWQ---NEGAGA 223
DB 122 PMEMLKIQLDAGRLAVHHQGSASAPTSRSYTTGASTHRRPRTATLAWELLRTQGLAG 181

QY 224 FYRSYTTTQTMNVPQAIHFMTYEFLOE--HFNPPQRRYNPSH--VLSGACAGAVAAATTP 281
DB 182 LYRGLGATLLRDIPFSIIYFFPLFANLNGLFNELAGKAFSAHFSVGCAGVAGIAVATP 241

QY 282 LDVCKTLTNTQESLALNSHITGHTMASAFRTVYQVGVTAFFRGVQARVYQIPSTA 340
DB 242 LDVLKTRITQTLKKGLEDWY-----SGITDCARKLWIEGSPAFMKGAGCRALVIAPLFG 296
```

QY 341 IAWSVY 346
|| ||
Db 297 IAQGVY 302

Search completed: August 27, 2002, 04:16:45
Job time: 1237 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:55:57 ; Search time 138.54 Seconds
(without alignments)
154.737 Million cell updates/sec

Title: US-09-870-113-4
Perfect score: 1036
Sequence: 1 MELEGRGAGVAGGPAAGPG.....GAAGCVATLLHDAAMNPAEG 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	786	75.9	318 22	AA1980
2	624	60.2	289 21	AA1981
3	624	60.2	289 22	AA1982
4	516.5	49.9	299 22	AA1983
5	515	49.7	331 22	AA1984
6	486.5	47.0	677 22	AA1985
7	477	46.0	268 22	AA1986
8	465	44.9	366 22	AA1987
9	440.5	42.5	155 21	AA1988
10	363	35.0	272 21	AA1989
11	295	28.5	379 22	AA1990

12	241.5	23.3	126	22	ABB51123
13	223	21.5	331	21	AAG22077
14	223	21.5	331	21	AAG43092
15	211	20.4	289	18	AAW17054
16	211	20.4	290	21	AAG22078
17	211	20.4	290	21	AAG43093
18	167	16.1	278	21	AAG22079
19	167	16.1	278	21	AAG43094
20	165.5	16.0	59	22	AA1989
21	154.5	14.9	81	21	AA1990
22	150	14.5	377	22	AA1991
23	145	14.0	374	22	AA1992
24	145	14.0	374	22	AA1993
25	144.5	13.9	360	22	AA1994
26	144.5	13.9	399	22	AA1995
27	143.5	13.9	449	22	AA1996
28	142	13.7	413	22	AA1997
29	138	13.3	329	22	AA1998
30	138	13.3	329	22	AA1999
31	137	13.2	125	22	AA2000
32	136.5	13.2	304	21	AA1980
33	136	13.1	305	22	AA1981
34	135.5	13.1	285	21	AA1982
35	135.5	13.1	320	21	AA1983
36	135.5	13.1	320	22	AA1984
37	135.5	13.1	320	22	AA1985
38	135.5	13.1	335	21	AA1986
39	135.5	13.1	363	21	AA1987
40	131.5	12.7	297	22	AA1988
41	131.5	12.7	317	22	AA1989
42	131.5	12.7	320	22	AA1990
43	131	12.6	310	22	AA1991
44	131	12.6	311	21	AA1992
45	131	12.6	311	22	AA1993

ALIGNMENTS

RESULT 1	AA1980
ID	AA1980 standard; Protein; 318 AA.
AC	AA1980
DT	22-OCT-2001 (first entry)
DE	Human polypeptide SEQ ID NO 6436.
KW	Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; leukaemia.
OS	Homo sapiens.
PN	WO200153312-A1.
PD	26-JUL-2001.
PF	26-DEC-2000; 2000WO-US34263.
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.


```
XX PA (INCYTE GENOMICS INC.
XX PI
XX PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
XX PI Baughn MR, Azimzal Y, Lu DAM, Au-Young J, Patterson C;
XX DR WPI: 2001-041424/05.
XX DR N-PSDB; AAF27733.
XX PS
XX PT Isolated polypeptide with a human transport protein sequence is useful
XX PT for the diagnosis, prevention and treatment of disorders associated
XX PT with the immune, reproductive and cardiovascular systems -
XX CC Claim 2; Page 133-134; 165pp; English.
XX CC The present invention provides the protein and coding sequences for 43
XX CC novel human transport proteins (designated TPPTS). These can be used in
XX CC the diagnosis and treatment of transport, metabolic, neurological,
XX CC reproductive, cardiovascular and immune disorders, and cell proliferative
XX CC disorders such as cancer.
XX SQ Sequence 289 AA;
XX
Query Match 60.2%; Score 624; DB 22; Length 289;
Best Local Similarity 100.0%; Pred. NO. 2.1e-54;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 MVAGAVAGILEHCVMYPIDCVKTRMSLOPPDPAARYRNVLALWRIIRTEGLWRPGRCLN 135
DB 1 mvagavagilehcvmypidcvktrmqlsqdpaaaryrnvlalwriirteglwrpgrln 60
QY 136 VTATGAGPAHALYFACYEKIKKTLSDVTHPGGNSHANGAGCVATLLHDAAMPAAE 192
DB 61 vtatgagpahalyfacyekikktlsdvthpggnsanganiagagcvatllhdaampae 117
RESULT 4
AAM79039
ID AAM79039 standard; Protein; 299 AA.
XX AC AAM79039;
XX DT
XX DT 06-NOV-2001 (first entry)
XX DE Human protein SEQ ID NO 1701.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorder; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200157190-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US04098.
XX PR 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PR 20-JUN-2000; 2000US-0598075.
XX PR 19-JUL-2000; 2000US-0620325.
XX PR 01-SEP-2000; 2000US-0654936.
XX PR 15-SEP-2000; 2000US-0663561.
XX PR 20-OCT-2000; 2000US-0693325.
XX PR 30-NOV-2000; 2000US-0728422.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
```

```
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI: 2001-476283/51.
XX DR N-PSDB; AAK52172.
XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX PT useful in diagnosis and gene therapy -
XX PS Claim 20; Page 4043; 622lpp; English.
XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX CC (AAM80020) are omitted as the relevant pages from the sequence listing
XX CC were missing at the time of publication.
XX SQ Sequence 299 AA;
XX
Query Match 49.9%; Score 516.5; DB 22; Length 299;
Best Local Similarity 55.7%; Pred. NO. 1.3e-43;
Matches 107; Conservative 19; Mismatches 39; Indels 27; Gaps 4;
QY 1 MELEGRGAGGVAGGPAAGRSPGESALLDGLQGVGRGAGGAGACRPVQDPDPSG 60
DB 28 melr---sgsv--gsqavarrmdgds-----rdgggg-----kdatgs 60
QY 61 PDYEALPAGATVTTMVGAVAGVAGILEHCVMYPIDCVKTRMSLOPPDPAARYRNVLALWR 120
DB 61 edyenlptsasvsthtmagamagilehsvmypdvsvktrmqlsqdpkagqtsvygalkk 120
QY 121 IIRTEGLWRPGRCLNVTATGAGPAHALYFACYEKIKKTLSDVTHPGGNSHANGAGCVA 180
DB 121 imrtgefwrplrgvnmimgagpahamyfacyenmkrtlndvfhqhgnshianglagma 180
QY 181 TLLHDAAMPAAE 192
DB 181 tllhdavmpae 192
RESULT 5
AAB60658
ID AAB60658 standard; Protein; 331 AA.
XX AC AAB60658;
XX DT 04-MAY-2001 (first entry)
XX DE Human mitochondrial solute carrier protein hMSC-o.
XX KW Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
XX KW preparation; detection.
XX OS Homo sapiens.
XX PN CN1269409-A.
XX PD 11-OCT-2000.
XX PF 17-MAR-2000; 2000CN-0114958.
XX PR 17-MAR-2000; 2000CN-0114958.
XX PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
```

XX Zhang X, Gao X, Xiao H;
 XX WPI: 2001-050544/07.
 DR N-PSDB; AAF59920.
 XX
 PT New human mitochondrion solute carrier protein and its nucleic acid -
 XX
 PS Claim 4; Page 20; 2lpp; Chinese.
 XX
 CC The invention relates to a novel human mitochondrial solute carrier
 CC protein, hMSC-o (AAB0658), and cDNA encoding it (AAF59920). hMSC-o is
 CC expressed in normal human hypothalamus tissue. The invention also relates
 CC to the preparation of hMSC-o proteins and nucleic acids, and the
 CC detection of hMSC-o proteins and nucleic acids in a sample. The present
 CC sequence represents hMSC-o.
 XX
 SQ Sequence 331 AA;

Query Match 49.7%; Score 515; DB 22; Length 331;
 Best Local Similarity 61.8%; Pred. No. 2,1e-43;
 Matches 97; Conservative 18; Mismatches 32; Indels 10; Gaps 1;
 QY 36 GVGAGGAGGACACPPVPRQDPGPDYEALPAGATVTTTHVAGAVAGILEHCVMPIDC 95
 Db 3 gdsrdggg-----kdatgsedyenlptsasvsthtagmagillehsvmpvds 52
 QY 96 VKTRMQLQDPDPAARYRNVLALRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKL 155
 Db 53 vktrmqslpdpkagqtslygalkkrmtegrwprlgvnmimgagpahamyfacyenn 112
 QY 156 KKTLSVVIHPGNSHIANGAGCVATLLHDAAMNPAE 192
 Db 113 krtldvfhqgnshtangagsmatllhdavmpae 149

RESULT 6
 AAM39748
 ID AAM39748 standard; Protein; 677 AA.
 XX
 AC AAM39748;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #239.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US08656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-611725/70.
 XX
 CC Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 PS Claim 20; Page 189; 765pp; English.
 XX

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX
 SQ Sequence 677 AA;

Query Match 47.0%; Score 486.5; DB 22; Length 677;
 Best Local Similarity 53.1%; Pred. No. 3.6e-40;
 Matches 102; Conservative 20; Mismatches 43; Indels 27; Gaps 4;
 QY 1 MELEGRGAGGAVAGPAAAGPGRSPGESALLDGLQGVGRGAGGACACRPVPRQDPDSG 60
 Db 13 melr---sgsv--gsqavarmdgd-----rdggg-----kdatgs 45
 QY 61 PVEALPAGATVTTTHVAGAVAGILEHCVMPIDCVKTRMQLQDPDPAARYRNVLALWR 120
 Db 46 edyenlptsasvsthtagmagillehsvmpvdsvktrmqslpspsqspvsiygalkk 105
 QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKKTLSVVIHPGNSHIANGAGCVA 180
 Db 106 lmrtegrwprlgvnmimgagpahamyfacyennkrtldvfhqgnshtangagsma 165
 QY 181 TLLHDAAMNPAE 192
 Db 166 tlldhdaavmpae 177

RESULT 7
 AAM39719
 ID AAM39719 standard; Protein; 268 AA.
 XX
 AC AAM39719;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2864.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AAI58875.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Example 4; SEQ ID NO 2864; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
XX specification.
XX Sequence 268 AA;
SQ

Query Match 46.0%; Score 477; DB 22; Length 268;
Best Local Similarity 94.7%; Pred. No. 1e-39;
Matches 90; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 98 TRMQSLQPPAARYNVLEALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKK 157
Db 2 trmhsllsqdpagrynvlevlwrirrtglwrpmlgnvtatgawpahalyfacyekllk 61
QY 158 TLDSDVHPGNSHIANGAGCVATLLHDAAMNPAE 192
Db 62 tldsdvhpqgnshiangaagcvatilhdampae 96

RESULT 8
AAM00938
ID AAM00938 standard; Protein; 366 AA.
XX
XX AAM00938;
XX
XX 01-OCT-2001 (first entry)
XX Human bone marrow protein, SEQ ID NO: 414.
DE
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX
OS Homo sapiens.
XX
XX WO200153453-A2.
XX
XX 26-JUL-2001.
PD
XX 23-DEC-2000; 2000WO-US34960.

XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX WPI: 2001-488707/53.
DR N-PSDB; AAI90057.
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX Claim 10; Page 504-505; 648pp; English.
XX The present sequence is one of 251 novel human polypeptides encoded
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX Sequence 366 AA;
SQ

Query Match 44.9%; Score 465; DB 22; Length 366;
Best Local Similarity 70.0%; Pred. No. 2.4e-38;
Matches 84; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

QY 73 TTHVAGAVAGILHCVMTPIDCVKTRMSQLQPPAARYNVLEALWRIIRTEGLWRP 132
Db 74 sthmtagmagilehsvmypvdsvktrmqslqppdkagysiygalkkimrtgfwrpir 133
QY 133 GLNVTATGAGPAHALYFACYEKLKTLSDVHPGNSHIANGAGCVATLLHDAAMNPAE 192
Db 134 gnvnmimgagpahamyfacyenmkrtindvfhqgnshlangiagsmatllhdavmnpae 193

RESULT 9
AAB42966
ID AAB42966 standard; Protein; 155 AA.
XX
XX AAB42966;
XX
XX 08-FEB-2001 (first entry)
XX Human ORFX OR2730 polypeptide sequence SEQ ID NO:5460.
DE
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipapillary; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200058473-A2.
XX
XX
PD 05-OCT-2000.
XX
XX
PF 31-MAR-2000; 2000WO-US08521.
XX
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
PA
PI Shimkets RA, Leach M;
XX
XX
DR WPI; 2000-602362/57.
DR N-PSDB; AAC77175.
XX
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX
PS Claim 11; Page 4641; 5507pp; English.
XX
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX
SQ Sequence 155 AA;

Query Match 42.5%; Score 440.5; DB 21; Length 155;
Best Local Similarity 52.3%; Pred. No. 2.5e-36;
Matches 91; Conservative 20; Mismatches 36; Indels 27; Gaps 4;

QY 1 MELBGRGAGGAGGAPAGPGRSPCESALLDGLQGVGRGAGGACRPPVRQDPDSG 60
Db 1 melr---sgsv---gsqavarmdgds-----rdgggg-----kdatgs 33
QY 61 PDYEALPAGATVTHMVAGVAGILEHCVMPIDCVKTRMQSLQDPDPAARYNVLEALWR 120
Db 34 edyenlptsasvsthtmagamagilehsvmvpdvsktrmqslspdpkagvtsiygalkk 93
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKTLSDVTHPGNSHIANG 174
Db 94 imqtegtwrlgrgvnmimgagpahamyfacyenmkrtlndvfhqgnsnhiang 147

RESULT 10
AAB42980
ID AAB42980 standard; Protein; 272 AA.
XX
XX
AC AAB42980;
XX
XX
DT 08-FEB-2001 (first entry)
XX
XX
DE Human ORFX ORF2744 polypeptide sequence SEQ ID NO:5488.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency (SCID); malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08521.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC77189.

XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 4662-4663; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance

CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC disorders (e.g. corneal graft neovascularisation and diabetic
CC retinopathy), neurological disorders (e.g. Huntington's chorea,
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. ABA8185 to
CC ABA83193 and ABA850300 represent sequences used in the exemplification of
CC the present invention.

XX.
SQ Sequence 126 AA;

Query Match 23.3%; Score 241.5; DB 22; Length 126;
Best Local Similarity 48.6%; Pred. No. 1.8e-16;
Matches 54; Conservative 11; Mismatches 23; Indels 23; Gaps 3;

Qy 36 GVGRCAGGGEAGACRPPVRQDPDSGPDYALPAGATVTHMVAGAVAGILEHCVMYPIDC 95

Db 36 gdsrdgggg-----kdtsgedyenlptsasvsthtagmagilehsvmypyds 85

Qy 96 VKTRMQSLQPPAARYNRNVLEALMRIITEGLWRPMRGLNVTATGAGPAHA 146

Db 86 vktrmqslspdpkadytsiygalkkimrte-----asg-gpcea 123

RESULT 13

AAG22077

ID AAG22077 standard; Protein: 331 AA.

XX AC AAG22077;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 24869.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX XX 06-SEP-2000.

XX XX 25-FEB-2000; 2000EP-0301439.

XX XX 25-FEB-1999; 99US-0121825.

XX XX 05-MAR-1999; 99US-0123180.

XX XX 09-MAR-1999; 99US-0123548.

XX XX 23-MAR-1999; 99US-0125788.

XX XX 25-MAR-1999; 99US-0126264.

XX XX 29-MAR-1999; 99US-0126785.

XX XX 01-APR-1999; 99US-0127462.

XX XX 06-APR-1999; 99US-0128234.

XX XX 08-APR-1999; 99US-0128714.

XX XX 16-APR-1999; 99US-0129845.

XX XX 19-APR-1999; 99US-0130077.

XX XX 21-APR-1999; 99US-0130449.

XX XX 23-APR-1999; 99US-0130510.

XX XX 23-APR-1999; 99US-0130891.

XX XX 28-APR-1999; 99US-0131449.

XX XX 30-APR-1999; 99US-0132048.

XX XX 30-APR-1999; 99US-0132407.

XX XX 04-MAY-1999; 99US-0132484.

XX XX 05-MAY-1999; 99US-0132485.

XX XX 06-MAY-1999; 99US-0132486.

XX XX 07-MAY-1999; 99US-0132487.

XX XX 11-MAY-1999; 99US-0132863.

XX XX 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

```
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.5%; Score 223; DB 21; Length 331;
Best Local Similarity 39.9%; Pred. No. 4e-14;
Matches 59; Conservative 22; Mismatches 51; Indels 16; Gaps 5;

QY 52 PVRODPDGDYEALPAGAVTT-----HMVAGAVAGILEHCVMYPIDCVKTRMQSLOPD 106
Db 16 pipppdfhp---aaiypagnttlkfwqlmvgagsvehmanfpvdtvktmqlrsc 72
QY 107 PAARYNVLEALWRIIRTEGLWRPMRLNVTATGAGPAHALYFACYEKLKLTLSVDIHPG 166
Db 73 p-ikpigrqfarslktktdpsalrygiwanglgagbahavfyfeyevskfls-----g 126
QY 167 G--NSHIANGAAGCVATLLHDAAMNPAE 192
Db 127 gpnnsaahaisgvfatissdavftpm 154

RESULT 14
AAG43092
ID AAG43092 standard; Protein; 331 AA.
XX AC AAG43092;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 53820.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PF 25-FEB-1999; 99US-0121825.
XX PF 05-MAR-1999; 99US-0123180.
XX PF 09-MAR-1999; 99US-0123548.
XX PF 23-MAR-1999; 99US-0125788.
XX PF 25-MAR-1999; 99US-0126264.
XX PF 29-MAR-1999; 99US-0126785.
XX PF 01-APR-1999; 99US-0127462.
XX PF 06-APR-1999; 99US-0128234.
XX PF 16-APR-1999; 99US-0128714.
XX PF 19-APR-1999; 99US-0129845.
XX PF 19-APR-1999; 99US-0130077.
XX PF 21-APR-1999; 99US-0130449.
XX PF 23-APR-1999; 99US-0130510.
XX PF 23-APR-1999; 99US-0130891.
XX PF 30-APR-1999; 99US-0131449.
XX PF 30-APR-1999; 99US-0132048.
XX PF 04-MAY-1999; 99US-0132407.
XX PF 05-MAY-1999; 99US-0132484.
XX PF 06-MAY-1999; 99US-0132485.
XX PF 06-MAY-1999; 99US-0132486.
XX PF 07-MAY-1999; 99US-0132487.
XX PF 07-MAY-1999; 99US-0132863.
```

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 21.5%; Score 223; DB 21; Length 331;
Best Local Similarity 39.9%; Pred. No. 4e-14;
Matches 59; Conservative 22; Mismatches 51; Indels 16; Gaps 5;
QY 52 PVRQDPDSDPYDALPAGATVTT-----HMVAGAVAGILEHCVMYPIDCVKTRMQSLQPD 106
Db 16 pibqpdpfhp---aaiivpaqnttlkfwlmvagsiagsvehmamfpvdtvktmqlrsc 72
QY 107 PAARYRNVLALWRIIRTEGLWRPMBGLNVTATGAGPAHALYFACYEKLKLTSLSDVIHPG 166
Db 73 p-ikpigrqafsrslktgpsalyrgiwamglagapahavysfyevskkfls-----g 126
QY 167 G--NSHIANGAAGCVATLLHDAAMNPAE 192
Db 127 gnpnnsaahaisgvfatissdavftpmd 154
RESULT 15
AAW17054
ID AAW17054 standard; Protein; 289 AA.
AC AAW17054;
XX
XX 05-AUG-1997 (first entry)
XX Blackcurrant RIB7 polypeptide.
XX
XX Blackcurrant; fruit-specific promoter; RIB7; transgenic plant.
XX OS Ribes nigrum strain Ben Alder.
XX PN W09717452-AL.
XX PD 15-MAY-1997.
XX
XX 04-NOV-1996; 96WO-EP04807.
XX
XX 03-NOV-1995; 95GB-0022558.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Brennan RM, Taylor MA, Woodhead MR;
XX PI
XX WPI: 1997-281041/25.
XX N-PSDB; AAT68957.
XX New isolated promoters from blackcurrant fruit - used for driving
XX fruit-specific expression of DNA sequences in transgenic
XX blackcurrant and other non-climacteric fruit
XX Example 4; Page 35-36; 66pp; English.
XX
XX RIB polypeptides (AAW17050-54) are encoded by cDNA clones (AAT68953-
XX 57) that exhibit differential expression in blackcurrant fruit
XX during the ripening period of fruit development. RIB7 shows
XX 62% similarity to yeast MRS4, a yeast mitochondrial RNA splicing
XX protein. RIB7 is expressed almost entirely in fruit. The
XX promoter region (AAT68952) of the RIB7 gene can be used as a
XX fruit-specific promoter.
XX

SQ Sequence 289 AA;
Query Match 20.4%; Score 211; DB 18; Length 289;
Best Local Similarity 38.5%; Pred. No. 5.5e-13;
Matches 45; Conservative 26; Mismatches 42; Indels 4; Gaps 2;
QY 76 MVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNVLALWRIIRTEGLWRPMBGLN 135
Db 1 miagsiagsiehmamypvdtlkttriqai-gscsaqsaglrqalgsilkvesgpaglyrgig 59
QY 136 VTATGAGPAHALYFACYEKLKLTSLSDVIHPGGNSHIANGAAGCVATLLHDAAMNPAE 192
Db 60 amglgagpahavysfyevmcketfs---hgdpsngahavsgvfatvasdavitpmd 113
Search completed: August 27, 2002, 03:55:58
Job time: 4910 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:58:49 ; Search time 48.11 Seconds
(without alignments)
97.987 Million cell updates/sec

Title: US-09-870-113-4

Perfect score: 1036

Sequence: 1 MELEGRGAGGVAGGPAAGPG.....GAAGCVATILLHDAMNPAEG 193

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCRUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	212	20.5	328	4	US-09-068-140A-15
2	211	20.4	289	4	US-09-068-140A-10
3	135.5	13.1	320	2	US-08-933-750C-12
4	135.5	13.1	320	4	US-09-234-613-12
5	131	12.6	351	2	US-08-933-750C-19
6	131	12.6	351	4	US-09-234-613-19
7	124.5	12.0	311	2	US-08-775-009-32
8	120	11.6	311	2	US-08-775-009-33
9	115	11.1	447	4	US-09-160-119-4
10	115	11.1	674	4	US-09-160-119-2
11	112.5	10.9	306	5	PCT-US94-09799-1
12	106.5	10.3	469	4	US-09-188-930-339
13	102.5	9.9	432	2	US-08-937-466-4
14	102.5	9.9	432	2	US-09-172-528-4
15	102.5	9.9	432	3	US-09-318-199-4
16	102.5	9.9	432	4	US-09-503-579-4
17	98.5	9.5	308	2	US-08-937-466-2
18	98.5	9.5	308	2	US-09-172-528-2
19	98.5	9.5	308	3	US-09-318-199-2
20	98.5	9.5	308	4	US-09-503-579-2
21	97	9.4	180	4	US-08-937-263B-8
22	94.5	9.1	312	4	US-09-142-565-2
23	94	9.1	180	2	US-08-791-495-9
24	94	9.1	312	4	US-09-188-930-142
25	93.5	9.0	256	2	US-08-937-466-6
26	93.5	9.0	256	2	US-09-172-528-6
27	93.5	9.0	256	3	US-09-318-199-6

Sequence 6, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 10, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 41, Appl
Sequence 41, Appl
Sequence 105, App

28 93.5 9.0 256 4 US-09-503-579-6
29 92.5 8.9 307 2 US-08-807-861A-56
30 92.5 8.9 307 3 US-09-210-681-56
31 92.5 8.9 307 3 US-08-946-719A-56
32 92.5 8.9 309 1 US-08-518-878B-51
33 92.5 8.9 309 2 US-08-807-861A-51
34 92.5 8.9 309 3 US-08-470-868A-51
35 92.5 8.9 309 3 US-09-210-681-51
36 92.5 8.9 309 3 US-08-946-719A-51
37 91 8.8 298 3 US-08-961-871-10
38 89.5 8.6 103 1 US-08-209-747-4
39 89.5 8.6 103 1 US-08-458-298-4
40 89.5 8.6 299 1 US-08-518-878B-56
41 89.5 8.6 299 2 US-08-470-868A-56
42 88.5 8.5 1332 1 US-07-609-716-41
43 88.5 8.5 1332 4 US-08-475-411A-41
44 88.5 8.5 1332 4 US-08-478-029A-41
45 87.5 8.4 146 1 US-07-609-716-105

ALIGNMENTS

RESULT 1
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-15

Db	66	CLLYCNCVLEPLYLCPNGARCATWFDQPRFTCTMDAFVKIVRHGRTTLW----
QY	138	ATGAGPAHALYFYCYEKLKKTLL-----SDVIHPGNSHTANGAAGCVATLLHDAAMNPA
Db	123	LVMTVPATAYFTAYDQLKCAFLCGRALTSOLYAP-----NVAGALARLGTVTVISPL
QY	192	E 192
Db	175	E 175

```

RESULT      7
US-08-775-009-32
; Sequence 32, Application US/08775009
; Patent No. 5935783
; GENERAL INFORMATION:
; APPLICANT: Gong, Weilong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
; APPLICANT: Roe, Bruce
; TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
; TITLE OF INVENTION: Velloiocardiofacial Syndrome Minimal Critical Region
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5935783ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,009

```

```

CLASSIFICATION: 433
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yatko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CH-0681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-009-32

```

```

Query Match      12.0%; Score 124.5; DB 2; Length 311;
Best Local Similarity 31.4%; Pred. No. 3,7e-05;
Matches 32; Conservative 22; Mismatches 43; Indels 5; Gaps 2;

QY    65 ALPA-GATVTH-----MVAGAVAGILHCVMYPIDCVKTRMQSLQPPAARYRNVLALW 119
      ||| | || :||: |||: |: ||::|: | |||: | |||: |
Db     13 AAPASGAKLTHPEKAILAGLAGGIEICITFPTEYYKTQLDERSHPHYRGIGDCVR 72
      ||| | || :||: |||: |: ||::|: | |||: | |||: |

QY    120 RIITEGLWRPMPRLNVTATCAGPAHAHALYFACYEKLKLTSD 161
      |: |: | |||: | | | | | | | | | | | |
Db     73 QTVRSHGVGLYRGLSLLYGSIPTKAARVFQMFEFLSNHMRD 114
      |||: | |||: | | | | | | | | | | | |

RESULT      8
US-08-775-009-33
; Sequence 33, Application US/08775009

```

; Patent No. 5935783
; GENERAL INFORMATION:
; APPLICANT: Gong, Weillong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
; APPLICANT: Roe, Bruce
; TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
; TITLE OF INVENTION: Velicardiofacial Syndrome Minimal Critical Region
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5935783ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,009
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yanko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CH-0681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-775-009-33

Query Match 11.6%; Score 120; DB 2; Length 311;
Best Local Similarity 29.6%; Pred. No. 0.00011;
Matches 29; Conservative 23; Mismatches 42; Indels 4; Gaps 1;
QY 68 AGATVTH-----MVAGVAGILEHCVMPIDCVKTRMOSLQDPDPAARYNVLEALWRIIR 123
DB 17 SGKALTHPGKAILAGLAGGAGTEICITPTTEYVKTQLQDLDERANPPRYRGIGDCVQRQTVR 76
QY 124 TEGLRWPRMGLNVTATGAGPAHALYFACYEKLKKTLSLD 161
DB 77 SHGVGLYRGUSSLLYGSIPKAAVRFGMFEFLSNHMRD 114
RESULT 9
US-09-160-119-4
; Sequence 4, Application US/09160119A
; Patent No. 6316219
; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRIL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160.119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 447
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-4

Query Match 11.1%; Score 115; DB 4; Length 447;
Best Local Similarity 30.1%; Pred. No. 0.00057;
Matches 34; Conservative 18; Mismatches 39; Indels 22; Gaps 4;
QY 79 GAVAGILEHCVMPIDCVKTRMOSLQDPDPA-----ARYRVNLEALWRIIRTEGLWRPMRGL 134
DB 107 GSVAGAVGATAVPIDLKTRMQNQSTGSGFVGMELMYKNSFDCFKVLYRGEFFGLYRGL 166
QY 135 NVTATGAGPAHALYFACYEKLKKTLSDDVI-----HPGGN-----SHIANGAAG 177
DB 167 LPQLLGVAPEKA-----IKLTVNDVFRDKFMHKGDSVPLAAETLAGGCAG 211

RESULT 10
US-09-160-119-2
; Sequence 2, Application US/09160119A
; Patent No. 6316219
; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRIL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160.119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 674
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-2

Query Match 11.1%; Score 115; DB 4; Length 674;
Best Local Similarity 30.1%; Pred. No. 0.00098;
Matches 34; Conservative 18; Mismatches 39; Indels 22; Gaps 4;
QY 79 GAVAGILEHCVMPIDCVKTRMOSLQDPDPA-----ARYRVNLEALWRIIRTEGLWRPMRGL 134
DB 334 GSVAGAVGATAVPIDLKTRMQNQSTGSGFVGMELMYKNSFDCFKVLYRGEFFGLYRGL 393
QY 135 NVTATGAGPAHALYFACYEKLKKTLSDDVI-----HPGGN-----SHIANGAAG 177
DB 394 LPQLLGVAPEKA-----IKLTVNDVFRDKFMHKGDSVPLAAETLAGGCAG 438

RESULT 11
PCT-US94-09799-1
; Sequence 1, Application PC/TUS9409799
; GENERAL INFORMATION:
; APPLICANT: Kagan, David
; TITLE OF INVENTION: Method and Composition for Weight Reduction
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Watov & Kipnes, P.C.
; STREET: 186 Princeton-Hightstown Rd, PO Box 247
; CITY: Princeton Junction
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08550

```

RESULT 12
US-09-188-930-339
; Sequence 339, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolating
; FILE OF INVENTION: and Methods For Their
; FILE REFERENCE: 11000.101ic1
; CURRENT APPLICATION NUMBER: US/09/188-930-339
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 1.0
; SEQ ID NO 339
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-339

```

RESULT 14
US-09-172-528-4


```

RESULT 15
US-09-318-199-4
; Sequence 4, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA

```

Search completed: August 27, 2002, 03:58:50
Job time: 4892 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:57:40 ; Search time 80.04 Seconds
(without alignments)
231.700 Million cell updates/sec

Title: US-09-870-113-4
Perfect score: 1036
Sequence: 1 MELEGRGAGVGAGGAAAGPG.....GAAGCVATLLHDAAMNPAEG 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	338.5	32.7	312	2 T26089	hypothetical prote
2	252	24.3	303	2 T39149	probable RNA splic
3	251.5	24.3	314	2 S5179	mRNA splice defec
4	224.5	21.7	781	2 A86205	hypothetical prote
5	223.5	21.6	304	2 S13533	mRNA splice defec
6	223	21.5	331	2 T00582	probable mitochond
7	168	15.2	307	2 S00949	probable phosphate
8	150	14.5	310	2 T20229	hypothetical prote
9	143	13.8	358	2 T45934	hypothetical prote
10	142	13.7	348	2 D84798	probable mitochond
11	134.5	13.0	436	2 J01459	Btl protein precur
12	132.5	12.8	291	2 T37992	probable tricarbox
13	131	12.6	351	2 T43493	hypothetical prote
14	129	12.5	313	2 D84613	hypothetical prote
15	128.5	12.4	332	2 T47703	Ca-dependent solu
16	128	12.4	352	2 T01729	mitochondrial solu
17	124.5	12.0	296	2 B96830	hypothetical prote
18	124.5	12.0	311	2 G01789	citrate transporte
19	124	12.0	301	1 S31935	ADP,ATP carrier pr
20	124	12.0	330	2 T09362	hypothetical prote
21	124	12.0	902	1 S54495	probable carrier p
22	123	11.9	367	2 G96770	hypothetical prote
23	122	11.8	392	2 T05350	adenylate transloc
24	121.5	11.7	310	2 S69050	probable membrane
25	121.5	11.7	313	2 T05577	uncoupling protein
26	121	11.7	650	2 T32897	hypothetical prote
27	120.5	11.6	322	2 S71116	probable carrier p
28	120	11.6	311	2 A46595	tricarboxylate tra
29	119.5	11.5	311	2 G86383	probable mitochond

30	118.5	11.4	353	2 T51393	probable mitochond
31	118.5	11.4	363	2 T49281	mitochondrial phos
32	118	11.4	330	2 T29640	mitochondrial carr
33	116	11.2	296	2 T23170	hypothetical prote
34	115.5	11.1	702	2 T16533	hypothetical prote
35	115	11.1	368	2 S54524	hypothetical prote
36	114.5	11.1	314	2 A56650	2-oxoglutarate car
37	114.5	11.1	314	2 A36305	2-oxoglutarate/mal
38	114.5	11.1	374	2 T24162	hypothetical prote
39	114	11.0	312	2 H88567	protein KILH3.3 [i
40	114	11.0	366	2 T01169	phosphate transport
41	113	10.9	299	2 S44554	citrate transport
42	112.5	10.9	307	2 G01858	uncoupling protein
43	112	10.8	335	2 T50393	probable mitochond
44	112	10.8	363	2 T47240	amino acid transpo
45	112	10.8	381	2 T51158	hypothetical prote

ALIGNMENTS

RESULT 1

T26089
hypothetical protein W02B12.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T26089
R:Swinsburne, J.; Ainscough, R.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z20149
A:Accession: T26089
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-312 <WIL>
A:Cross-references: EMBL:Z66521; PIDN:CAA91399.1; GSPDB:GN00020; CESP:W02B12.9
A:Experimental source: clone W02B12
C:Genetics:
A:Gene: CESP:W02B12.9
A:Map position: 2
A:Introns: 18/3; 251/3; 286/3
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match	32.7%	Score 338.5;	DB 2;	Length 312;
Best Local Similarity	48.0%;	Pred. NO. 8.8e-21;		
Matches 73;	Conservative 23;	Mismatches 37;	Indels 19;	Gaps 4;
QY	42	GGGEAGACPPVRQDPDSDGPDYEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQ	101	
Db	3	GGGE-----DEYESLPT-HSVPVHLTAGALAGAVEHCVMFPDPSVKTRMQ	46	
QY	102	SLQPPDPAARYNVLEALWRIIRTEGLWRPMRGLNTATGAGPAHALYFACYEKLKLTSD	161	
Db	47	SLCPCEPKCPTPVHSLMSIVKREGWLRLRGVNAAGSMPAHALYFTVYERKMKGYLTG	106	
QY	162	VIHPGNSH-IANGAAGCVATLLHDAAMNPAE	192	
Db	107	--NSAGHSNTLAYAGSVVATLIHDAIMNPAE	136	

RESULT 2

T39149
probable RNA splicing proteinmitochondrial carrier protein - fission yeast (Schizosacc
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39149
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21748
A:Accession: T39149
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-303 <OLI>

A;Cross-references: EMBL:X87371; NID:G954542; PID:CAA60822.1; PTD:G954563
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C;Genetics:
A;Gene: SGD:MRS3
A;Cross-references: SGD:S0003669; MIPS:YJL133w
A;Map position: 10L
A;Genome: nuclear

C:Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology
C:Keywords: mitochondrial inner membrane; mitochondrion; pre-mRNA splicing
F:30-119/Domain: ADP, ATP carrier protein repeat homology <ACP1>
F:127-211/Domain: ADP, ATP carrier protein repeat homology <ACP2>
F:216-311/Domain: ADP, ATP carrier protein repeat homology <ACP3>
F:127-211/Domain: ADP, ATP carrier protein repeat homology <ACP2>
F:216-311/Domain: ADP, ATP carrier protein repeat homology <ACP3>

Query Match	24.3%	Score 251.5;	DB 2;	Length 314;
Best Local Similarity	41.0%;	Pred. No. 1.4e-13;		
Matches	57;	Conservative 18;	Mismatches 63;	Indels 1;
Gaps				
Qy	52	PVRODPSGPDYEALPAGATVTTTHMVAVAGVAGILEHCVMYPIDCVKTPTMOSLOPDPAARY	111	
	1;			

Qy	112	RNVLEALWRIIRTEGLWRPWRGLNVTATGAPAHALFYACYEKLKKYPLSDVHPPGGNSHI	171
		: : : : : :	
Dd	72	NNLMSQISHISTSEGTIALMKGVQSIVILGAPAHAVFGTYEFCKKNLIDSSDTQTHHPF	131
		: : : : : :	
Qy	172	ANGAAGCVATILLDHAMNP	190
		: : : : : :	
Dd	132	KTAISGCATTASDALMNP	150
		: : : : : :	

RESULT 4

A86205
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A86205
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizarr, L.

[illegible]

A:Cross-references: EMBL:X92441; NID:g1050762; PIDN:CAA63185.1; PID:g1050774
C:Genetics:
A:Map position: 15R
A:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Note: YOR222w
C:Keywords: duplication; mitochondrion; transmembrane protein
F:9-107/Domain: ADP,ATP carrier protein repeat homology <ACPI>
F:115-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:208-300/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 16.2% Score 168; DB 2; Length 307;
Best Local Similarity 33.9%; Pred. No. 1.2e-06;
Matches 42; Conservative 20; Mismatches 54; Indels 8; Gaps 1;

Qy 77 VAGAVAGILEHCVMYPIDCVKTRMSQLQPDPAA-----RYRNVLLEALWRIIRTEGLW 128
 :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 17 ISGAVAGISELTVMYPLDVVKTRFQLEVTFPTAAAGVQOVERYGVIDCLKKIIVKKEGFS 76

Qy 129 RPMGLNVNTAGAPAHAFVCYEKKUKKLTSVDVIHPGGNSHIANGAACGVATLLHDAAM 188
 | ||| | | | | | : | ||| : | ||| : | ||| : | ||| : | ||| : |
Db 77 RLYRGISSPMLMEAPKRATKFACNDQYQKIFKNLFNTNETTKTQKISIAAGSAGMTEAAVI 136

Qy 189 NPAE 192
 | |
Db 137 VPFE 140

RESULT 8
T20229
hypothetical protein C54G10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T20229
R:Matthews, L.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19240
A:Accession: T20229
A>Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: DNA
A:Residues: 1-310 <WIL>
A:Cross-references: EMBL:Z75532; PIDN:CAA99811.1; GSPDB:GN00023; CESP:C54G10.4
A:Experimental source: clone C54G10
C:Genetics:
A:Gene: CESP:C54G10.4
A:Map position: 5
A:Introns: 14/1; 63/1; 155/2; 256/2; 299/3
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

R; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F. submitted to the Protein Sequence Database, January 2000

A: Reference number: Z23017

A: Accession: T45934

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-358 <MON>

A: Cross-references: EMBL:AL132960

A: Experimental source: cultivar Columbia; BAC clone F5K20

C: Genetics:

A: Map position: 3

A: Introns: 98/3; 157/3; 207/3; 234/3; 258/2; 281/1

A: Note: F5K20.240

C: Superfamily: ADP-ATP carrier protein; ADP-ATP carrier protein repeat homology

[illegible]

RESULT 8
T20229
hypothetical protein C54G10.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T20229
R:Matthews, L.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19240
A:Accession: T20229
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-310 <MIL>
A:Cross-references: EMBL:Z75532; PIDN:CAA99811.1; GSPDB:GN00023; CESP:C54G10
A:Experimental source: clone C54G10
C:Genetics:
A:Gene: CESP:C54G10.4
A:Map position: 5
A:Introns: 14/1; 63/1; 155/2; 256/2; 299/3
C:Superfamily: ADP.ATP carrier protein; ADP.ATP carrier protein repeat homolog

Query Match	14.5%	Score 150;	DB 2;	Length 310;
Best Local Similarity	28.9%;	Pred. No. 3.6e-05;		
Matches	44;	Conservative 23;	Mismatches 55;	Indels 30; Gaps 6;
Qy	43	GGEAGACRPVPRODPSGDYEAALPAGATVTHMVAGAVAGILSHCVMPIDCVKTRMQS	102	
Db	80	GVHGTCRQ--MEDPDS-----ITSFVGCAAGMAQSVAAPATERIKLLIQ-	124	
Qy	103	LQDDPA-ARYRNVLAEALWRIIRTEGLWRPRMGLNVATGAGPAHALAFACYEKLKK----	157	
Db	125	IODDKAHTKNGPIDATKQLLRTHGLKSIRFGTFLATVARADAPGVCFASVENMARSMCK	184	
Qy	158	-----TLSDVIHPGGNSHANGAACGVATLLH	184	
Db	185	DGETSTLS-----SGOLLFAGGTAGMLSWLFN	211	

RESULT 9
T45934
hypothetical protein F5K20_240 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000
C:Accession: T45934

	Query Match	13.8%;	Score 143;	DB 2;	Length 358;
	Best Local Similarity	28.8%;	Pred. No. 0.00016;		
	Matches	36;	Conservative 23;	Mismatches 56;	Gaps 10;
Qy	68	AGATVTTHTM	VAGAVAGTLEHCVMY	PIDCVTRMQSLQPDPAA	RNRNVLEALWRIIRTEGL 127
		: : : : : : : : : : : :			
Db	172	AGVDISIVHFVS	GGLAGTAASATYPDL	VLRLUSA-----QG	VGHAFRTICREEGI 232
		: : : : : : : : : : : :			
Qy	128	WRPMRGENVLTATG	AGPAHALYFYACYEKLK	TLSDVTHPGGNSHIANGA	ACGVATLLHDA 187
		: : : : : : : : :			
Db	223	LGLYKGLCATLLG	VGPSLAISFAYETP-KTF	WLSHRPNDNSNAVSLGC	SGLSIVSSTA 281
		: : : : : : : : : : : :			
Qy	188	MNPAE	192		
	:				
Db	282	TFPD	286		

RESULT 10
D84798
probable mitochondrial carrier protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: D84798
R;Lin, X.; Kaul, S.; Rounslev, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanNaken, S.E.; Umayam, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
A;Reference number: AB84420; MUID:20083487

A:Residues: 1-348 <STO>
A:Cross-references: GB:AE002093; NID:g4895195; PIDN:AAD3782.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g37890
A:Map position: 2
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: mitochondrion

	Query Match	13.7%	Score 142;	DB 2;	Length 348;
	Best Local Similarity	32.8%;	Pred. No. 0.00019;		
	Matches 40;	Conservative 20;	Mismatches 46;	Indels 16;	Gaps 4
QY	75	HWGAVAGILEHCVMYPTDCVKTRMQSLDPPAARYNVLEALWRIIRTEGLWRPMRGL	134		
Db	151	HFVSGGLAGTAAATATYPLDVRTLAQR--NAITYOGIEHTFTTICREEGILGYLKG	208		
QY	135	NVTATGAGPAHALYFACYEKLK-----KTLSDVTHFGGNSHANGAAGCVATL---LH	184		
Db	209	GATILGVGPSIAINFAAYESMKLEWHSHRPNDSDIV---VSLVSGGLAGAVSSSTDDKLF	264		

Qy	185	DA	186
Db	265	DA	266

Db 4 YDPKPGNNNTKFOVAVAGSVGLVTRALISPFVDVIRKQF-LQHERLSRSDPSAKYHGI 62
QY 115 LEALWRIIRTEG---LWRPMGLNVTATGAGPAHALYFACYEKLKTLSDVIRTEGPGG---- 167
Db 63 LQASRQILQEGPTAFWKGHVPAQILSIGYAVOFLSF-----EMLTVLVHRGSVYDA 115

QY 168 ---NSH-IANGAGCVATL 182
Db 116 REFSVHFVCGGLAACMATL 134

RESULT 5

BT1_MAIZE ID BT1_MAIZE STANDARD; PRT; 436 AA.
AC P29518;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brittle-1 protein, chloroplast precursor.
GN BT1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9305685; PubMed=1668652;
RA Sullivan T.D., Strelow L.I., Illingworth C.A., Phillips R.L.,
RA Nelson O.E. Jr.;
RT "Analysis of maize brittle-1 alleles and a defective
RT suppressor-mutator-induced mutable allele.";
RL Plant Cell 3:1337-1348(1991).
CC -1- FUNCTION: COULD PLAY A ROLE IN AMYLOPLAST MEMBRANE TRANSPORT.
CC -1- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
CC -1- TISSUE SPECIFICITY: ENDOSPERM OF DEVELOPING KERNELS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M79333; AAA33438.1; -.
DR PIR; JQ1459; JQ1459.
DR MaizeDB; 47578; -.
DR InterPro; IPR002067; Mit_carrier.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Transit peptide; Chloroplast; Amyloplast; Transmembrane.
FT TRANSIT 1 75 CHLOROPLAST (POTENTIAL).
FT CHAIN 76 436 BRITTLE-1 PROTEIN.
FT TRANSMEM 229 247 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
SQ SEQUENCE 436 AA; 46627 MW; 9600C03F603E9DAE CRC64;

Query Match 13.0%; Score 134.5; DB 1; Length 436;
Best Local Similarity 34.3%; Pred. No. 0.00058;
Matches 42; Conservative 16; Mismatches 52; Indels 13; Gaps 4;
QY 67 PAGATVTTNHWAGAVAGILEHCVMYPIDCVKTRMQSLQDPAPAAARNVLEALWRIIRTEG 126
Db 224 PPKPIPTPLVAGLAGFASTLCTYPEMLIKTRV-TIEKDV---YDNVAHFVKILRDEG 279
QY 127 LWRPMGLNVTATGAGPAHALYFACYEKLKTLSDV-I--HPGNSHTANGAAGCVATLLH 184
Db 280 PSELYRGLTSLIGVVPYAAACNFYAYETLKRRLYRRATGRRPGAD-----VGPVATLLI 332

QY 185 DAA 187
Db 333 GSA 335

RESULT 6

ODC_HUMAN ID ODC_HUMAN STANDARD; PRT; 299 AA.
AC Q9BQ78;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
GN Mitochondrial 2-oxodicarboxylate carrier.
GN SLC25A21 OR ODC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RX TISSUE=Liver;
RX MEDLINE=21269385; PubMed=11083877;
RA Fiermonte G., Dolce V., Palmieri L., Ventura M., Runswick M.J.,
RA Palmieri F., Walker J.E.;
RT "Identification of the human mitochondrial oxodicarboxylate carrier:
RT bacterial expression, reconstitution, functional characterization,
RT tissue distribution and chromosomal location.";
RL J. Biol. Chem. 276:8225-8230(2001).
CC -1- FUNCTION: Transports C5-C7 oxodicarboxylates across the inner
CC membranes of mitochondria. Can transport 2-oxoadipate, 2-
CC oxoglutarate, adipate, glutarate, and to a lesser extent,
CC pimelate, 2-oxopimelate, 2-aminoadipate, oxaloacetate, and
CC citrate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ278148; CAC27562.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 17 37 POTENTIAL.
FT TRANSMEM 62 82 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 211 231 POTENTIAL.
SQ SEQUENCE 299 AA; 33303 MW; 69A259400328AE19 CRC64;

Query Match 12.9%; Score 134; DB 1; Length 299;
Best Local Similarity 31.7%; Pred. No. 0.00043;
Matches 38; Conservative 26; Mismatches 48; Indels 8; Gaps 4;
QY 76 MVAGAVAGILEHCVMYPIDCVKTR--MQSLQDPAPAAARNVLEALWRIIRTEGLWRPMRG 133
Db 17 IVAGSGNLVEICMLPHLDVVKTRFQIQRCATDPNS-YKSLVDSFRMIFQMEGLFGYKG 75
QY 134 LNTATGAGPAHALYFACYEKLKTLSDV-IHPGNSHTANGAAGCVATLLHDAAMPAPAE 192
Db 76 ILPPIAETPKRAVKFTTFQYKLLGVLSLSPALTPAIAGLSGLTEAIV---VNPFE 131

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF317711; AAG60687.1; -
 DR EMBL; AF151827; AAD34064.1; -
 DR EMBL; AL133584; CAB63728.1; -
 DR EMBL; AK026060; BAB15341.1; -
 DR EMBL; BC001398; AAH01398.1; -
 DR EMBL; BC009330; AAH09330.1; -
 DR EMBL; AF119864; AAF69618.1; -
 DR InterPro; IPR001993; Mitoch_carrier.
 DR InterPro; IPR002067; Mit_carrier.
 DR Pfam; PF00153; mito_carr; 3.
 DR PROSITE; PS00926; MITOCARRIER.
 DR PROSITE; PS00215; MITOCH_CARRIER; 1.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Alternative splicing; Polymorphism.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 364 382 POTENTIAL.
 FT VARSPLIC 64 72 LPSSLSQSTG -> W (IN ISOFORM 2).
 FT VARIANT 247 247 L -> F.
 FT FTID=VAR_012756.
 FT CONFLICT 266 266 G -> R (IN REF. 6).
 FT SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;

Query Match 12.3%; Score 127; DB 1; Length 359;
 Best Local Similarity 23.8%; Pred. No. 0.0019;
 Matches 45; Conservative 19; Mismatches 51; Indels 74; Gaps 5;

QY 67 PAGATVTHWVAGAVAGILEHCVMYPIDCVKTRMSLQPDPA----- 109
 DB 6 PAGISPLQOMVASGTGAVVTVSLFMTPLDVVKVRLQSRPMSASELMPSSRLWSLSTKL 65
 QY 110 -----RYRNVLEALWRIIRTEG---LWR 129
 DB 66 SLSQSTGKLLYCNGLVLEPLVLCNPGARCATWFDPTRTGTMDAFVKIVRHEGRTLW- 124
 QY 130 PMRGLNVATGAGPAHALYFACYEKLKTL-----SDVIHPGGNSHIANGAGCVATLL 183
 DB 125 --SGLPATLVMTVPATAYFTAYDQKAFKLCGRALTSOLYAP-----WVAGALARLG 174
 QY 184 HDAAMNPAE 192
 DB 175 TVTVISPLE 183

RESULT 8
 TXTP_BOVIN
 ID TXTP_BOVIN STANDARD; PRT; 311 AA.
 AC P79110;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tricarboxylate transport protein, mitochondrial precursor (Citrate
 DE transport protein) (CTP) (Tricarboxylate carrier protein).
 GN SLC20A3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.

RESULT 7
 CG69_HUMAN
 ID CG69_HUMAN STANDARD; PRT; 359 AA.
 AC Q9BJJ4; Q9JUF6; Q9Y379; Q9P182;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mitochondrial carrier protein CGI-69 (PRO2163).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SUBCELLULAR LOCATION, AND
 RP TISSUE SPECIFICITY.
 RC TISSUE=Liver;
 RX PubMed=11139402;
 RA Yu X.X., Lewin D.A., Zhong A., Brush J., Schow P.W., Sherwood S.W.,
 RA Pan G., Adams S.H.;
 RT "Overexpression of the human 2-oxoglutarate carrier lowers
 RT mitochondrial membrane potential in HEK-293 cells: contrast with the
 RT unique cold-induced mitochondrial carrier CGI-69.";
 RL Biochem. J. 353:369-375(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20272150; PubMed=10810093;
 RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
 RT "Identification of novel human genes evolutionarily conserved in
 RT Caenorhabditis elegans by comparative proteomics.";
 RL Genome Res. 10:703-713(2000).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Testis;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansgore W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain, and Colon;
 RA Strausberg R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 15-359 FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
 RA He F.;
 RT "Functional prediction of the coding sequences of 79 new genes deduced
 RT by analysis of cDNA clones from human fetal liver.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC 1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC 2- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC 3- ALTERNATIVE PRODUCTS: 2 isoforms; 1/CGI-69L (shown here) and 2;
 CC are produced by alternative splicing.
 CC 4- TISSUE SPECIFICITY: Expressed in many tissues. Abundant in testis
 CC and kidney.
 CC 5- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RX MEDLINE=21154914; PubMed=11230163;
 RA Wu Q., Zhang T., Cheng J.-F., Kim Y., Grimwood J., Schmutz J.,
 RA Dickson M., Noonan J.P., Zhang M.Q., Myers R.M., Maniatis T.;
 RT "Comparative DNA sequence analysis of mouse and human protocadherin
 RT gene clusters.";
 RL Genome Res. 11:389-404(2001).
 CC -!- FUNCTION: Ornithine transport across inner mitochondrial membrane,
 CC from the cytoplasm to the matrix (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF332005; AAK26320.1; -
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; 2.
 DR Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 KW TRANSMEM 5 25 POTENTIAL.
 FT TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 110 130 POTENTIAL.
 FT TRANSMEM 168 188 POTENTIAL.
 FT TRANSMEM 210 230 POTENTIAL.
 FT TRANSMEM 237 257 POTENTIAL.
 SQ SEQUENCE 301 AA; 32580 MW; F22760BADE31DE21 CRC64;

 Query Match 11.8%; Score 122; DB 1; Length 301;
 Best Local Similarity 33.3%; Pred. No. 0.0041;
 Matches 28; Conservative 16; Mismatches 38; Indels 2; Gaps 1;
 QY 76 MVAGAVAGILEHCVMYPIDCVKRMQSLQDPAAARNVLEALWRIITGLWPRMGLN 135
 Db 213 MLSGGVAGICLWLWVFFVDICKSRQVL--SMYKQAGFGITLLSVVRNEGIVALYSGLK 270
 QY 136 VTATGAGPAHALYFACYEKLKLT 159
 Db 271 ATMIRAIPANGALFVAVYEYSRKMM 294
 RESULT 11
 ODC1_YEAST
 ID ODC1_YEAST STANDARD; PRT; 310 AA.
 AC OQ3028;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DE 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mitochondrial 2-oxodicarboxylate carrier 1.
 GN ODC1 OR YPL134C OR LPI11C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97313271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
 RA Araujo R., Aparicio A., Barrell B., Badcock K., Benes V., Botstein D.,
 RA Bowman S., Bruckner M., Carpenter J., Cherry J.M., Chung E.,
 RA Churcher C., Coster F., Davis K., Davis R.W., Dietrich F.S.,

RA Delius H., DiPaolo T., Dubois E., Dusterhoft A., Duncan M., Floeth M.,
 RA Fortin N., Friesen J.D., Fritz C., Goffeau A., Hall J., Hebling U.,
 RA Heumann K., Hilbert H., Hillier L., Hunnicke-Smith S., Hymen R., D.,
 RA Johnston M., Kaiman S., Kleine K., Komp C., Kurdi O., Lashkari F.,
 RA Lew H., Lin A., Lin D., Louis E.J., Marathe R., Messenguy F.,
 RA Mewes H.W., Mirtipati S., Moestl D., Muller-Auer S., Namath A.,
 RA Nurnelle U., Oefner P., Pearson D., Petel F.X., Pohl T.M.,
 RA Purnelle D., Schafer M., Scharfe M., Scherens B., Schramm S.,
 RA Schroeder M., Sdicu A.M., Tettelin H., Urrestarazu L.A., Ushinsky S.,
 RA Vierendeels F., Viissers S., Voss H., Walsh S.V., Wambutt R., Wang Y.,
 RA Wedler E., Wedler H., Winnett E., Zhong W.W., Zollner A., Vo D.H.,
 RA Hani J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
 RL Nature 387:103-105(1997).
 RN [2]
 RP CHARACTERIZATION.
 RX PubMed=11013234;
 RA Walker J.E.;
 RT "Identification in Saccharomyces cerevisiae of two isoforms of a novel
 RT mitochondrial transporter for 2-oxoadipate and 2-oxoglutarate.";
 RL J. Biol. Chem. 276:1916-1922(2001).
 CC -!- FUNCTION: Transports C5-C7 oxodicarboxylates across the inner
 CC membranes of mitochondria. Can transport 2-oxoadipate, 2-
 CC oxoglutarate, adipate, glutarate, 2-oxopimelate, oxaloacetate,
 CC citrate and malate. The main physiological role is probably to
 CC supply 2-oxoadipate and 2-oxoglutarate from the mitochondrial
 CC matrix to the cytosol where they are used in the biosynthesis of
 CC lysine and glutamate, respectively, and in lysine catabolism.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U43703; AAB68225.1; -
 DR SGD; S0006055; ODC1.
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 FT TRANSMEM 9 29 POTENTIAL.
 FT TRANSMEM 126 146 POTENTIAL.
 FT TRANSMEM 179 199 POTENTIAL.
 FT TRANSMEM 219 239 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 SQ SEQUENCE 310 AA; 34206 MW; 8EB7D6231BC2BCEE CRC64;

 Query Match 11.7%; Score 121.5; DB 1; Length 310;
 Best Local Similarity 30.0%; Pred. No. 0.0046;
 Matches 36; Conservative 19; Mismatches 46; Indels 19; Gaps 4;
 QY 78 AGAVAGILEHCVMYPIDCVKTRMQ---SLQDPDPA---RYRNVLEALWRIIRTEG 126
 Db 17 AGATAGVELLYVYPLDVVTRMQLQVTKGHPAAVAAKAADVHTGVMDCLTIVKKEG 76
 QY 127 LWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVTIHGNSHI-----ANGAGCV 179
 Db 77 FSHLYKGTITSPILMEAPKRAIKFSGNDTF-QTYKKFFPTPGEMTOKIATYSGASAGV 135
 RESULT 12
 SFCL_YEAST
 ID SFCL_YEAST STANDARD; PRT; 322 AA.
 AC P33303;


```
RESULT 14
MFT_HUMAN
ID MFT_HUMAN STANDARD; PRT; 315 AA.
AC Q9H2D1;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial folate transporter/carrier.
GN MFT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538421; PubMed=10978331;
RA Titus S.A.; Moran R.G.;
RT "Retrovirally mediated complementation of the glyB phenotype. Cloning
RT of a human gene encoding the carrier for entry of folates into
RT mitochondria."
RL J. Biol. Chem. 275:36811-36817(2000).
CC -!- FUNCTION: Transport folate across the inner membranes of
CC mitochondria.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF283645; AAC37834.1; -.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PRINTS: PR00926; MITOCARRIER.
DR PROSITE: PS00215; MITOCH_CARRIER; 1.
DR Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 89 106 POTENTIAL.
FT TRANSMEM 227 243 POTENTIAL.
FT TRANSMEM 281 300 POTENTIAL.
SQ SEQUENCE 315 AA; 35388 MW; 60EBC0D61951EB6E CRC64;

Query Match 11.5%; Score 119.5; DB 1; Length 315;
Best Local Similarity 27.5%; Pred. No. 0.0068;
Matches 41; Conservative 23; Mismatches 54; Indels 31; Gaps 7;

QY 68 AGATVTH-----MVAGAVAGILEHCVMYPIDCVKTRMQ-----SIQPPAARYNRVLEA 117
DB 13 AWSTVFRHVRVYNELAGVSGVLSNLALHPLDLVLRFAVSDGLRPP-----KYNGLHC 68
QY 118 LWRIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLK--KTLSDVIHPGNGNSHIANGA 175
DB 69 LTTINKLGLRGLYGVTPNINAGLSWGLYFFFNAYKSYKTEGRAHLEATEYLVSA 128
QY 176 -AG-----CVAT-----LLHDAAMN 189
DB 129 EAGATMLCITNPWTKTRMLQYDAVYN 157

RESULT 15
ADT_ANOGA
ID ADT_ANOGA STANDARD; PRT; 301 AA.
AC Q27238;
```

```
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ADP/ATP carrier protein (ADP/ATP translocase) (Adenine nucleotide
DE translocator) (ANT).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=G3;
RX MEDLINE=94348635; PubMed=8069414;
RA Beard C.B.; Crews-Oyen A.E.; Kumar V.K.; Collins F.H.;
RT "A cDNA encoding an ADP/ATP carrier from the mosquito Anopheles
RT gambiae."
RL Insect Mol. Biol. 3:35-40(1994).
CC -!- FUNCTION: CATALYZES THE EXCHANGE OF ADP AND ATP ACROSS THE
CC MITOCHONDRIAL INNER MEMBRANE.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L11618; AAB04104.1; -.
DR EMBL: L11617; AAB04105.1; -.
DR InterPro: IPR002067; Mit_carrier.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PRINTS: PR00926; MITOCARRIER.
DR PROSITE: PS00215; MITOCH_CARRIER; 3.
DR Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 14 31 1 (POTENTIAL).
FT TRANSMEM 75 93 2 (POTENTIAL).
FT TRANSMEM 119 136 3 (POTENTIAL).
FT TRANSMEM 178 197 4 (POTENTIAL).
FT TRANSMEM 216 233 5 (POTENTIAL).
FT TRANSMEM 275 293 6 (POTENTIAL).
SQ SEQUENCE 301 AA; 32863 MW; 4CC9E17C9F8DA08B CRC64;

Query Match 11.4%; Score 118.5; DB 1; Length 301;
Best Local Similarity 29.2%; Pred. No. 0.0078;
Matches 31; Conservative 19; Mismatches 51; Indels 5; Gaps 3;

QY 78 AGAVAGILEHCVMYPIDCVKTRM-QSLOQDPAAR-YRNVLEALWRIIRTEGLWRPMRGLN 135
DB 121 SGAGAGATSLCFVYPLDFARTRLGADVGPAGAGERENGLLDCLKTKVKSDDGIIGLRGN 180
QY 136 VTATGAGPAHALYFACYEKLKRTKLTSDVIHPGNGNSHIANGAAGCVAT 181
DB 181 VSVQGIIVYRAAYFGCFDTAKGMLPD---PKNTSIFVSWAIQVVT 223

Search completed: August 27, 2002, 04:16:46
Job time: 1238 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:19:20 ; Search time 130.92 seconds
(without alignments)
255.026 Million cell updates/sec

Title: US-09-870-113-4

Perfect score: 1036

Sequence: 1 MELEGRGAGGVAGGPAAGPG.....GAAGCVATLLHDAAMNPAEG 193

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriaph.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1030	99.4	364	4 Q96A46	Q96A46 homo sapien
2	516.5	49.9	338	11 Q920G8	Q920G8 mus musculus
3	515.5	49.8	347	4 Q9NYZ2	Q9NYZ2 homo sapien
4	444.5	42.9	155	4 Q969S1	Q969S1 homo sapien
5	444.5	42.9	182	11 Q91ZY0	Q91ZY0 mus musculus
6	338.5	32.7	312	5 Q23125	Q23125 caenorhabdi
7	295	28.5	379	5 Q9VAY3	Q9VAY3 drosophila
8	291	28.1	380	5 Q9NHV6	Q9NHV6 onchocerca
9	267.5	25.8	303	5 Q94638	Q94638 onchocerca
10	265.5	25.6	301	5 Q94634	Q94634 onchocerca
11	252	24.3	303	3 Q14281	Q14281 schizosacch
12	224.5	21.7	781	10 Q9LMJ6	Q9LMJ6 arabidopsis
13	223	21.5	331	10 Q64731	Q64731 arabidopsis
14	211	20.4	289	10 Q82049	Q82049 ribes nigru
15	158.5	15.3	147	11 Q9D547	Q9D547 mus musculus
16	155.5	15.0	322	10 Q9AX03	Q9AX03 oryza sativ

17	150	14.5	310	5 Q18844	Q18844 caenorhabdi
18	144.5	13.9	320	11 Q921P8	Q921P8 mus musculus
19	144.5	13.9	399	5 Q9VKZ5	Q9VKZ5 drosophila
20	143.5	13.9	449	5 Q9VI05	Q9VI05 drosophila
21	143	13.8	358	10 Q9M333	Q9M333 arabidopsis
22	142	13.7	348	10 Q9SH98	Q9SH98 arabidopsis
23	142	13.7	413	5 Q9VWF9	Q9VWF9 drosophila
24	139	13.4	298	11 Q99JD3	Q99JD3 rattus norv
25	137.5	13.3	321	4 Q9BSK2	Q9BSK2 homo sapien
26	136.5	13.2	304	4 Q9NWX2	Q9NWX2 homo sapien
27	136.5	13.2	304	4 Q96DW6	Q96DW6 homo sapien
28	136.5	13.2	334	11 Q9CW38	Q9CW38 mus musculus
29	136	13.1	319	10 Q9MA27	Q9MA27 arabidopsis
30	135.5	13.1	320	4 Q9HC21	Q9HC21 homo sapien
31	134.5	13.0	345	11 Q9CYJ1	Q9CYJ1 mus musculus
32	134	12.9	299	4 Q9BQT8	Q9BQT8 homo sapien
33	133.5	12.9	322	10 Q9ARL9	Q9ARL9 hordeum vul
34	133	12.8	361	3 Q9P6L7	Q9P6L7 schizosacch
35	132.5	12.8	291	3 Q13844	Q13844 schizosacch
36	131.5	12.7	297	5 Q9VBN7	Q9VBN7 drosophila
37	131.5	12.7	317	5 Q9VGM3	Q9VGM3 drosophila
38	131	12.6	310	4 Q9NVN5	Q9NVN5 homo sapien
39	131	12.6	311	4 Q96CO1	Q96CO1 homo sapien
40	131	12.6	311	11 Q922G0	Q922G0 mus musculus
41	131	12.6	351	4 Q9Y379	Q9Y379 homo sapien
42	131	12.6	351	4 Q9UF66	Q9UF66 homo sapien
43	130	12.5	305	10 Q92WG1	Q92WG1 arabidopsis
44	129	12.5	313	10 Q9SUY5	Q9SUY5 arabidopsis
45	128.5	12.4	306	5 Q9V3T2	Q9V3T2 drosophila

ALIGNMENTS

RESULT 1

Q96A46	PRELIMINARY;	PRT;	364 AA.
ID Q96A46			
AC Q96A46;			
DT 01-DEC-2001 (TREMBLrel. 19, Created)			
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DE MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE			
DE CARRIER SPLICE VARIANT).			
GN HMRS3/4.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE-21195335; PubMed-11297739;			
RA Li F., Nikali K., Grogan J., Leibiger I., Leibiger B., Schweyen R.,			
RA Larsson C., Suomalainen A.;			
RT "Characterization of a novel human putative mitochondrial transporter			
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.";			
RL FEBS Lett. 494:79-84(2001).			
DR EMBL; AJ303077; CAC27996.1; -.			
DR EMBL; AF327402; AAK49519.1; -.			
SQ SEQUENCE 364 AA; 39271 MW; 22049C4F8DD7A577 CRC64;			

Query Match 99.4%; Score 1030; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.3e-82; Indels 0; Gaps 0;
Matches 192; Conservative 0; Mismatches 0;

QY 1	MELEGRGAGGVAGGPAAGPGRSPGESALLDGLQVGRGAGGAGACACRPPVRQDPDSG 60	
DB 1	MELEGRGAGGVAGGPAAGPGRSPGESALLDGLQVGRGAGGAGACACRPPVRQDPDSG 60	
QY 61	PDYEALPAGATVTTTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYNNVLEALWR 120	
DB 61	PDYEALPAGATVTTTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYNNVLEALWR 120	

QY 121 IIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVIHPGGNSHIANGAAGCVA 180
Db 121 IIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVIHPGGNSHIANGAAGCVA 180
QY 181 TLLHDAAMNPAE 192
Db 181 TLLHDAAMNPAE 192
RESULT 2
Q920G8 PRELIMINARY; PRT; 338 AA.
AC Q920G8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Q.-Z., Ruan Q.-G., Eckenrode S., Shi J.-D., Cruz P., Wang C.-Y.,
RA She J.-X.;
RT "A new gene which is highly expressed in NOD mice spleen."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF288621; AAL23859.1; -
SQ SEQUENCE 338 AA; 37310 MW; BB35B1F70C56A3FE CRC64;

Query Match 49.9%; Score 516.5; DB 11; Length 338;
Best Local Similarity 54.7%; Pred. NO. 2e-37;
Matches 105; Conservative 19; Mismatches 41; Indels 27; Gaps 3;

QY 1 MELEGRGAGGAGGPAAGPGRSPGESALLDGLQRGVGRGAGGAGACRPPVRQDPDSG 60
Db 1 MELRRGGVGNAAAG-----RRMDGDCRDG--GCGSKDAGS----- 33
QY 61 PDYEALPAGATVTHWAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARYNRVLEALWR 120
Db 34 EDYENLPTSASVSTHTAGAMAGILEHSVMPVDSVKTRMQSLSDPKRAQYTSIYGALKR 93
QY 121 IIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVIHPGGNSHIANGAAGCVA 180
Db 94 IMRTGEGWRPLRGLNVMWMMGAGPAHAMYFACYENKMKRTLNDVFSHQGNSHLANGVAGSMA 153
QY 181 TLLHDAAMNPAE 192
Db 154 TLLHDAAMNPAE 165

RESULT 3
Q9NYZ2 PRELIMINARY; PRT; 347 AA.
AC Q9NYZ2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HT015 PROTEIN.
GN HT015.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-HYPOTHALAMUS;
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human hypothalamus."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF223466; AAF64141.1; -
DR InterPro: IPR001993; Mitoch_carrier.

DR Pfam: PF00153; mito_carrier; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWNL.2.
SQ SEQUENCE 347 AA; 37828 MW; 700D861B230E001E CRC64;

Query Match 49.8%; Score 515.5; DB 4; Length 347;
Best Local Similarity 55.2%; Pred. NO. 2.5e-37;
Matches 106; Conservative 20; Mismatches 39; Indels 27; Gaps 4;

QY 1 MELEGRGAGGAGGPAAGPGRSPGESALLDGLQRGVGRGAGGAGACRPPVRQDPDSG 60
Db 1 MELR---SGSV--GSQAVARRMDGDS-----RDGGG-----KDATGS 33
QY 61 PDYEALPAGATVTHWAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARYNRVLEALWR 120
Db 34 EDYENLPTSASVSTHTAGAMAGILEHSVMPVDSVKTRMQSLSDPKRAQYTSIYGALKK 93
QY 121 IIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVIHPGGNSHIANGAAGCVA 180
Db 94 IMRTGEGWRPLRGLNVMWMMGAGPAHAMYFACYENKMKRTLNDVFSHQGNSHLANGSMA 153
QY 181 TLLHDAAMNPAE 192
Db 154 TLLHDAAMNPAE 165

RESULT 4
Q969S1 PRELIMINARY; PRT; 155 AA.
AC Q969S1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MITOCHONDRIA SOLUTE CARRIER PROTEIN (HYPOTHETICAL 16.8 KDA
DE PROTEIN).
GN MSCP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Q., Eckenrode S., Ruan Q., Wang C., Shi J., McIndoe R.A., She J.;
RT "Molecular cloning of a novel mitochondria solute carrier protein
RT (MSCP) gene from mouse and human and its down-regulation in mouse
RT spleen during the maturation of the immune system."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS, AND LEIOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY032628; AAK38154.1; -
DR EMBL: BC015013; AAH15013.1; -
KW Hypothetical protein.
SQ SEQUENCE 155 AA; 16832 MW; 8ACB98A483C8E6EF CRC64;

Query Match 42.9%; Score 444.5; DB 4; Length 155;
Best Local Similarity 52.9%; Pred. NO. 1.6e-31;
Matches 92; Conservative 19; Mismatches 36; Indels 27; Gaps 4;

QY 1 MELEGRGAGGAGGPAAGPGRSPGESALLDGLQRGVGRGAGGAGACRPPVRQDPDSG 60
Db 1 MELR---SGSV--GSQAVARRMDGDS-----RDGGG-----KDATGS 33
QY 61 PDYEALPAGATVTHWAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARYNRVLEALWR 120
Db 34 EDYENLPTSASVSTHTAGAMAGILEHSVMPVDSVKTRMQSLSDPKRAQYTSIYGALKK 93
QY 121 IIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVIHPGGNSHIANG 174
Db 94 IMRTGEGWRPLRGLNVMWMMGAGPAHAMYFACYENKMKRTLNDVFSHQGNSHLANG 147

```
RESULT 5
Q912Y0 PRELIMINARY: PRT; 182 AA.
ID Q912Y0
AC Q912Y0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE MITOCHONDRIAL CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Li Q., Eckenrode S., Wang C., Ruan Q., Shi J., McIndoe R.A., She J.;
RT "A novel mouse mitochondrial carrier protein gene is up-regulated from
RL young to adult NOD mice.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF361699; AAL27990.1; -.
SQ SEQUENCE 182 AA; 19638 MW; 18E2C5E801228693 CRC64;

Query Match 42.9%; Score 444.5; DB 11; Length 182;
Best Local Similarity 52.3%; Pred. No. 2e-31;
Matches 91; Conservative 18; Mismatches 38; Indels 27; Gaps 3;

Qy 1 MELEGRGAGVAGGPAAGPGRSPGESALLDGLQGVGRGAGGAGAGCPRPVRQDPDSG 60
Db 1 MELRRGGVGNQAAG-----RMDGDCRDG---CGSKDAGS----- 33

Qy 61 PDYEALPAGATVTHWAGAVAGIIEHCVMYPIDCVKTRMQSLQDPDPAARYRNVLALWR 120
Db 34 EDYENLPTASVSTHMTAGAGIIEHSIMYPVDSYKTRMQSLNPDPKARYTSIYGALKR 93
I: :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 IIRTEGLRPMRLGNVTATGAGPAHALYFACYEKLKTKTLDVIHPGGNSHIANG 174
Db 94 IMHTGCFWRPLRGLNMMGAGPAHAMFYACVYENKMKTLNDVFSHQNSHLANG 147

RESULT 6
Q23125 PRELIMINARY: PRT; 312 AA.
ID Q23125;
AC Q23125;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE W02B12.9 PROTEIN.
GN W02B12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swinburne J., Ainscough R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA none;
RX MEDLINE=99069613; PubMed=9851916;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z66521; CAA91399.1; -.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;

Query Match 32.7%; Score 338.5; DB 5; Length 312;
Best Local Similarity 48.0%; Pred. No. 7e-22;
Matches 73; Conservative 23; Mismatches 37; Indels 19; Gaps 4;

Qy 42 GGGGAGACRPPVRQDPDSDGYEALPAGATVTHWAGAVAGIIEHCVMYPIDCVKTRMQ 101
Db 3 GGGG-----DEYESLPT-HSPVPHLTAGALAGAVEHCVMFPFDSVKTRMQ 46
I: |||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Qy 102 SLQPDPAARYRNVLALWRIIRTEGLRPMRLGNVTATGAGPAHALYFACYEKLKTKTLD 161
Db 47 SLQPCPETKCPPTFPVHSLMSIVKREGWLRPLRGVNAAGSMPAHALYFTVYEKKMGKYL 106
I: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Qy 162 VTHPGNSH-IANGAGACVATLLHDAAMPAAE 192
Db 107 --NSAGHSNTLAYGASGVVATLLHDAIMPAAE 136
I: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

RESULT 7
Q9VAY3 PRELIMINARY: PRT; 379 AA.
ID Q9VAY3;
AC Q9VAY3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE CG4963 PROTEIN (GH09840P).
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.B.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., DeCher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwag C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
```

```

[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Erise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; AE003763; AAF56764.1; -.
DR EMBL; AY060268; AAL25307.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 28.5%; Score 295; DB 5; Length 379;
Best Local Similarity 47.3%; Pred. No. 5.6e-18;
Matches 62; Conservative 22; Mismatches 41; Indels 6; Gaps 3;

QY 62 DYELPAGATVTTTHMVAGAVAGILEHCVMPIDCVKTRMQSLQDPAPAAARNVLEALWRI 121
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
DB 5 DYESLPT-TSVGVMNTAGAGVLEHVMVPLDSVTRMQSL--SPPTKMNIVSTLRTM 61
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
QY 122 IRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTLSDVIHPGNSHTANGAAGCVAT 181
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
DB 62 ITREGLLRPIRGASAVVLGAGPAHSLYFRAYEMTKELTAKFTSVRNLYVISGA---VAT 118
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
QY 182 LLHDAAMNPAE 192
| : | | | : | : |
DB 119 LIHDAISSPTD 129
| : | | | : | : |

RESULT 8
Q9NH96 PRELIMINARY; PRT; 380 AA.
AC Q9NH96;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 41.8 KDA PROTEIN.
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Schmid K.J., Tautz D.;
RT "A screen for rapidly evolving genes from Drosophila.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217402; AAF73387.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 41844 MW; 4C9AA524B97F8C6C CRC64;

Query Match 28.1%; Score 291; DB 5; Length 380;

```

```

Best Local Similarity 46.6%; Pred. No. 1.3e-17;
Matches 61; Conservative 22; Mismatches 42; Indels 6; Gaps 3;

QY 62 DYELPAGATVTTTHMVAGAVAGILEHCVMPIDCVKTRMQSLQDPAPAAARNVLEALWRI 121
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
DB 5 DYESLPT-TSVGVMNTAGAGVLEHVMVPLDSVTRMQSL--SPPTKMNIVSTLRTM 61
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
QY 122 IRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTLSDVIHPGNSHTANGAAGCVAT 181
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
DB 62 ITREGLLRPIRGASAVVLGAGPAHSLYFRAYEMTKELTAKFTSVRNLYVISGA---VAT 118
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
QY 182 LLHDAAMNPAE 192
| : | | | : | : |
DB 119 LIHDAISSPTD 129
| : | | | : | : |

RESULT 9
Q94638 PRELIMINARY; PRT; 303 AA.
AC Q94638;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45998; AAB19037.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 303 AA; 33861 MW; 224BFA547E5D617C CRC64;

Query Match 25.8%; Score 267.5; DB 5; Length 303;
Best Local Similarity 37.5%; Pred. No. 1.1e-15;
Matches 54; Conservative 26; Mismatches 41; Indels 23; Gaps 2;

QY 49 CRPPVRQDPDSDYEAALPAGATVTTTHMVAGAVAGILEHCVMPIDCVKTRMQSLQDPDA 108
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
DB 12 CRWPV-----HLLAGSVAGLAELHCLMFPPDSVKTRLOSCLPCPE 50
| | | | | : | : | | | | | | | | | | | | | | | : | | : |
QY 109 ARYRNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTLSDVIHPGNG 168
| : | : | : | | | : | | | : | | | | | | | | | | | : | | : |
DB 51 TSCPTAMHSLMSVKEGLRLSLKGVNAVVLGTPAHAFYTYVYENSKAYLLN--NPRVS 108
| : | : | : | | | : | | | : | | | | | | | | | | | : | | : |
QY 169 SHIANGAAGCVATLLHDAAMNPAE 192
| : | : | : | | | : | | | : | | | | | | | | | | | : | | : |
DB 109 NSVSYAISGALATVIHDAVMNPAE 132
| : | : | : | | | : | | | : | | | | | | | | | | | : | | : |

RESULT 10
Q94634 PRELIMINARY; PRT; 301 AA.
AC Q94634;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca gibsoni.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6284;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RT mitochondrial solute carriers.";
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45997; AAB19036.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 301 AA; 34176 MW; COBA8D819FB8EA79 CRC64;

Query Match 25.6%; Score 265.5; DB 5; Length 301;
Best Local Similarity 37.5%; Pred. No. 1.7e-15;
Matches 54; Conservative 26; Mismatches 41; Indels 23; Gaps 2;

Qy 49 CRPPVQPDSPDYALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPPDA 108
Db 12 CRVPV-----HLLAGSVAGLAELHMFPPDFSVKTRLSLQSLCPCE 50

Qy 109 ARYRNVLALWRIIRTEGLWRPMLGNVTATGAPAHALYFACYEKLKTTLSVDVTHPGGN 168
Db 51 TSCPAMHSLMSWKREGLLSLKGVNAVVLGTIPAHAFYTVYENSXAYLLN--NPRVS 108

Qy 169 SHIANGAAGCVATLLHDAAMNPAE 192
Db 109 NSMSVAISGALATVTHDAVMNPAE 132

RESULT 11
O14281 PRELIMINARY; PRT; 303 AA.
AC O14281;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE MITOCHONDRIAL CARRIER C8C9.12C.
GN SPAC8C9.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Oliver K., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RL - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC - DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC - SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; 299168; CAB16300.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Hypothetical protein: Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
SQ SEQUENCE 303 AA; 32652 MW; 65975CDE18107AB5 CRC64;

Query Match 24.3%; Score 252; DB 3; Length 303;
Best Local Similarity 40.3%; Pred. No. 2.5e-14;
Matches 52; Conservative 25; Mismatches 48; Indels 4; Gaps 1;

Qy 62 DYELPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPPDAARYRNVLALWRI 121
Db 9 DYELPIGSPMYAHLLAGAFSGILEHSVMYPVDAIKTRMQMLNGVRSVSGNIVNSVIKI 68
```

```

Qy 122 IRTEGLWRPMLGNVTATGAPAHALYFACYEKLKTTLSVDVTHPGNSHIANGAAGCVAT 181
Db 69 SSTEGLWIRGSISSVIMGAGPSHAIYFVLEFFSK-----INASDPRPLASALAGACAI 124

Qy 182 LLHDAAMNP 190
Db 125 TISDAFMTP 133

RESULT 12
O9LMJ6 PRELIMINARY; PRT; 781 AA.
AC O9LMJ6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE F10K1.26 PROTEIN.
GN F10K1.26.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu S.X., Chan A., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
RA Sakano H., Toriumi M., Chung M., Goldsmith A., Liu A., Liu A.,
RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RL "The sequence of BAC F10K1 from Arabidopsis thaliana chromosome 1.";
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC067971; AAF82217.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 4.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 781 AA; 87081 MW; 9FB579B9BD746D1E CRC64;

Query Match 21.7%; Score 224.5; DB 10; Length 781;
Best Local Similarity 38.0%; Pred. No. 1.9e-11;
Matches 57; Conservative 23; Mismatches 55; Indels 15; Gaps 5;

Qy 52 PVRQPD-----SGPDY--EALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQ 104
Db 9 PKFQEPDLRQVSQTPDFKPEIAHDLGLKFWQFIAGSVGVEHMAFPVDTIKTHMQALR 68

Qy 105 PDPAARYRNVLALWRIIRTEGLWRPMLGNVTATGAPAHALYFACYEKLKTTLSVDVTH 164
Db 69 PCP-LKPVGIREAFRSIIQKEGFSALYRGIMWGLGAPAHAVYFSFVEVSKKYL---- 123

Qy 165 PGG--NSHIANGAAGCVATLLHDAAMNPAE 192
Db 124 -AGDONNSVAHMSGVFATISSDAVFTPM 152

RESULT 13
O64731 PRELIMINARY; PRT; 331 AA.
AC O64731;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE MITOCHONDRIAL CARRIER PROTEIN.
GN AT2G30160.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```


Db 73 EDYENLPTSASVSTHMTAGAMAGILEHSIMYPVDSVK 109

Search completed: August 27, 2002, 04:19:22
Job time: 1289 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:55:58 ; Search time 138.54 Seconds
(without alignments)
212.463 Million cell updates/sec

Title: US-09-870-113-8

Perfect score: 1406

Sequence: 1 MQSLQDPAAARYNVLALW.....VYEFKYLITKQFEWRAGK 265

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1406	100.0	289	21	AAB50383	Human uncoupling p
2	1406	100.0	289	22	AAB60113	Human transport pr
3	1406	100.0	318	22	AAM41505	Human polypeptide
4	1377	97.9	268	22	AAM93719	Human polypeptide
5	1013.5	72.1	272	21	AAB42980	Human ORFX ORF2744
6	999.5	71.1	366	22	AAM00938	Human bone marrow
7	916	65.1	174	22	AAU01989	Human secreted pro
8	910	64.7	176	21	AAB40634	Human ORFX ORF398
9	910	64.7	176	22	AAB98221	Human mitochondria
10	848.5	60.3	331	22	AAB60658	Human mitochondria
11	754.5	53.7	677	22	AAU29748	Novel human secret

12	739.5	52.6	199	22	AAM80023	Human protein S80
13	725.5	51.6	187	22	AAM93909	Human polypeptide,
14	725.5	51.6	187	22	AAB98222	Human mitochondria
15	725.5	51.6	187	22	AAB49665	Human mitochondria
16	717.5	51.0	187	22	AAM93934	Human polypeptide,
17	717.5	51.0	188	21	AAB32112	Human secreted pro
18	692	49.2	299	22	AAM79039	Human protein S80
19	657.5	46.8	379	22	ABB61130	Drosophila melanog
20	558	39.7	215	22	AAB49666	O. volvulus mitoch
21	541	38.5	194	21	AAB32111	Human secreted pro
22	515.5	36.7	181	22	AAU01988	Gene #25 human sec
23	429.5	30.5	289	18	AAW17054	Blackcurrant Rib7
24	416	29.6	278	21	AAG22079	Arabidopsis thalia
25	416	29.6	290	21	AAG22078	Arabidopsis thalia
26	416	29.6	331	21	AAG22077	Arabidopsis thalia
27	414	29.4	278	21	AAG43094	Arabidopsis thalia
28	414	29.4	290	21	AAG43093	Arabidopsis thalia
29	414	29.4	331	21	AAG43092	Arabidopsis thalia
30	412	29.3	111	21	AAB42968	Human ORFX ORF2732
31	343.5	24.4	149	21	AAB42964	Human ORFX ORF2728
32	309	22.0	74	22	ABG01167	Novel human diagno
33	276	19.6	155	21	AAB42966	Human ORFX ORF2730
34	266	18.9	413	22	ABB69108	Drosophila melanog
35	261.5	18.6	358	22	ABB71142	Drosophila melanog
36	251.5	17.9	351	21	AAB50378	Human uncoupling p
37	243.5	17.3	365	21	AAB42783	Human ORFX ORF2547
38	242.5	17.2	301	21	AAG52781	Arabidopsis thalia
39	242.5	17.2	301	21	AAG52783	Arabidopsis thalia
40	242.5	17.2	325	21	AAG52780	Arabidopsis thalia
41	242.5	17.2	325	21	AAG52782	Arabidopsis thalia
42	242.5	17.2	361	21	AAG52779	Arabidopsis thalia
43	237	16.9	449	22	ABB59263	Drosophila melanog
44	233	16.6	296	21	AAB43372	Human ORFX ORF3136
45	232	16.5	374	22	AAM40959	Human polypeptide

ALIGNMENTS

RESULT 1

AAB50383
ID AAB50383 standard; Protein; 289 AA.

XX
AC AAB50383;

XX
DT 12-MAR-2001 (first entry)

XX
DE Human uncoupling protein #6.

XX
KW Human; uncoupling protein; immunosuppressive; antiarthritic;

XX
KW antirheumatic; antiproliferative; cardiant; vasotropic;

XX
KW cerebroprotective; neuroprotective, antibacterial; ophthalmological;

XX
KW gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;

XX
KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;

XX
KW infertility.

XX
OS Homo sapiens.

XX
PN WO2000061614-A2.

XX
PD 19-OCT-2000.

XX
PF 06-APR-2000; 2000WO-US09534.

XX
PR 09-APR-1999; 99US-0128701.

XX
PR 08-JUL-1999; 99US-0142821.

XX
PR 18-AUG-1999; 99US-0149448.

XX
PR 12-NOV-1999; 99US-0164751.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;

XX

DR WPI; 2000-656322/63.
DR N-PSDB; AAC90457.
XX
XX Uncoupling proteins and nucleic acid sequences encoding them, useful
PT for detecting, preventing and treating proliferative, neurological,
PT immune system, cardiovascular and gastrointestinal disorders -
XX
XX
PS Claim 11; Page 323-324; 343pp; English.
XX
XX The present sequence is a human uncoupling protein. The nucleotide
CC sequences encoding the uncoupling proteins may be used for
CC the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
CC or research purposes. The uncoupling protein encoded by the nucleotide
CC sequences may be used to treat disorders such as neural, immune,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders, wounds, infectious diseases,
CC thrombosis, arthritis, and infertility.
XX
SQ Sequence 289 AA;

Query Match 100.0%; Score 1406; DB 21; Length 289;
Best Local Similarity 100.0%; Pred. No. 3.1e-153;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOSLPDPAARYNVLEALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKKTLL 60
DB 25 mqslqpdparynvlealwrlrtedglwprmglnvtatgagpahalyfacyekllkktl 84
QY 61 SDVIHPGNSHIANGAAGCVATLLHDAAMPAAEVVKQMOMNSPYHRVDCVRAVWQNE 120
DB 85 sdvihpgnshiangagcvatlilhdaampaevkvqrmqmnspyrhrtidcvravwqne 144
QY 121 GAGAFYSYTTQLTMNVPPFOAIHPMTVEFLQEHNPORRNPSSHVLSGACAGAVAAAT 180
DB 145 gagafysyttqltmnvppfoaihpmfveflqehfnprnpsshvlsagacagavaaaat 204
QY 181 TPLDVCCKTLNTQESLALNSHITGHITGMASAFRTVYQVGVTAYFRGVQARVIYQIPST 240
DB 205 tpldvccktlntqeslalnshitghitgmasafrtvyqvgvtayfrgvqarvlyqipst 264
QY 241 AIAMSVYEFFKYLITKRQEWRACK 265
DB 265 aiawsvyeffkylitkrqeewrack 289

RESULT 2
AAB60113
ID AAB60113 standard; Protein; 289 AA.
XX
AC AAB60113;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33.
XX
KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer.
OS Homo sapiens.
XX
XX WO200078953-A2.
PN
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US16668.
XX
XX 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.

XX (INCY-) INCYTE GENOMICS INC.
XX
XX Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI; 2001-041124/05.
DR N-PSDB; AAF27733.
XX
XX Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
XX Claim 2; Page 133-134; 165pp; English.
XX
XX The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
SQ Sequence 289 AA;

Query Match 100.0%; Score 1406; DB 22; Length 289;
Best Local Similarity 100.0%; Pred. No. 3.1e-153;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOSLPDPAARYNVLEALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKKTLL 60
DB 25 mqslqpdparynvlealwrlrtedglwprmglnvtatgagpahalyfacyekllkktl 84
QY 61 SDVIHPGNSHIANGAAGCVATLLHDAAMPAAEVVKQMOMNSPYHRVDCVRAVWQNE 120
DB 85 sdvihpgnshiangagcvatlilhdaampaevkvqrmqmnspyrhrtidcvravwqne 144
QY 121 GAGAFYSYTTQLTMNVPPFOAIHPMTVEFLQEHNPORRNPSSHVLSGACAGAVAAAT 180
DB 145 gagafysyttqltmnvppfoaihpmfveflqehfnprnpsshvlsagacagavaaaat 204
QY 181 TPLDVCCKTLNTQESLALNSHITGHITGMASAFRTVYQVGVTAYFRGVQARVIYQIPST 240
DB 205 tpldvccktlntqeslalnshitghitgmasafrtvyqvgvtayfrgvqarvlyqipst 264
QY 241 AIAMSVYEFFKYLITKRQEWRACK 265
DB 265 aiawsvyeffkylitkrqeewrack 289

RESULT 3
AAM41505
ID AAM41505 standard; Protein; 318 AA.
XX
AC AAM41505;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6436.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF

XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 03-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI60661.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 2; SEQ ID NO 6436; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
 CC immunosuppressant and cytotstatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 318 AA;
 SQ

Query Match 100.0%; Score 1406; DB 22; Length 318;
 Best Local Similarity 100.0%; Pred. No. 3 6e-153;
 Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQSLQDPDPAARYRNVLALWRIIRTEGLWRPMLNVTATGAGPAHALYFACYEKLKTL 60
 Db |||||
 QY 54 mqslqpdpaaaryrnvllealwriirteglwrlpmlnvtatgagpahalyfacyekllkkl 113
 Db |||||
 QY 61 SDVIHPGCGNSHIANGAGCVATLLHDAAMPNPAEVVKORMQYNSPYHRVTDVRAVWQNE 120
 Db |||||
 QY 114 sdvihpgnshiangaagcvatlhlhdaampaeavvkqrmqynspyrhrvtdcvravwqne 173
 QY 121 GAGAFYSYTTQTMNVFQAIHMTVEYFLEHFNPPORRYNPSHVLGACAGAVAAAT 180
 Db |||||
 QY 174 gagafysyttqltmnvfqaifhmfeyfqlfhnpprrynpshvlsigacagavaaat 233
 Db |||||
 QY 181 TPLDVCKTLNTQESLALNSHITGTMASAFRTVYQVGVTAYFRGVQARVIYQIPST 240
 Db |||||
 QY 234 tpldvcktlntqeslalnshigtmasaftrvyqvgvtayfrgvqarviyqipst 293
 QY 241 AIAWSYEFKYLITRKQEWRAK 265
 Db |||||
 QY 294 aiawsyeffkylitkrqewragk 318

RESULT 4
 AAM39719
 ID AAM39719 standard; Protein: 268 AA.
 XX

AC AAM39719;
 XX 22-OCT-2001 (first entry)
 DT XX Human polypeptide SEQ ID NO 2864.
 DE XX
 XX Human; nootropic; immunosuppressant; cytotstatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 OS WO200153312-A1.
 PN XX
 XX 26-JUL-2001.
 PD XX
 XX 26-DEC-2000; 2000WO-US34263.
 PF XX
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI58875.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 4; SEQ ID NO 2864; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
 CC immunosuppressant and cytotstatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 268 AA;
 SQ

Query Match 97.9%; Score 1377; DB 22; Length 268;
 Best Local Similarity 98.1%; Pred. No. 6.2e-150;
 Matches 260; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MQSLQDPDPAARYRNVLALWRIIRTEGLWRPMLNVTATGAGPAHALYFACYEKLKTL 60
 Db |||||
 QY 4 mslqspdparyrnvllealwriirteglwrlpmlnvtatgawpahalyfacyekllkkl 63
 QY 61 SDVIHPGCGNSHIANGAGCVATLLHDAAMPNPAEVVKORMQYNSPYHRVTDVRAVWQNE 120

Db 64 sdvhpqgnshlangaagcvatilhdaamnpaevvkgrmqmynsp/hrvtdcvravwqne 123
QY 121 GAGAFYRSYTTQLTMNVPFQAIFHMTYEFQLQEHFNPPRRYNPSSHVLSCACAGAVAAAT 180
Db 124 gagafysyrtqltmvnpfqahmfmyeflqehfnpprrynpsshvlscagagavaaat 183
QY 181 TPLDVCKTLNLTQESLALNSHTGHTGMSAFRTYVQGVGTAYPRGVOARVIYQIPST 240
Db 184 tpldvcktlntqeslalsnshthgthgmsaftrtyvqgvgtayprgvarviyqipst 243
QY 241 ATAWSVEFFKYLITKROEWRACK 265
Db 244 atawsveffkylitkrqeagrak 268
RESULT 5
ID AAB42980 standard; Protein; 272 AA.
XX AAB42980;
XX
XX
DT 08-FEB-2001 (first entry)
DE Human ORFX ORF2744 polypeptide sequence SEQ ID NO:5488.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antiposrotatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX Homo sapiens.
XX
PN WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shmkets RA, Leach M;
PI WPI; 2000-602362/57.
DR N-PSDB; AAC77189.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 4662-4663; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antiposrotatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;

CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 272 AA;
Query Match 72.1%; Score 1013.5; DB 21; Length 272;
Best Local Similarity 69.7%; Pred. No. 4.9e-108;
Matches 184; Conservative 39; Mismatches 40; Indels 1; Gaps 1;
QY 1 MOSLPDPAARYRNVLALWRIIRTEGLWRPWRGLNVTATGAGPAHALYFACYEKLUKTL 60
Db 7 mqslspdpkaqtsiygalkimrtgfwrlgrgvnmimgagpahamyfacyennmkrtl 66
QY 61 SDVIHPGGNSHTANGAGCVATLLHDAAMNPAFVVKQRMQYNPSYHRYTDCVRAVWQNE 120
Db 67 ndvfhqgnshlangagsmatllhdavmnpaevvkqrlqmynpsqrsalscirtvwrt 126
QY 121 GAGAFYRSYTTQLTMNVPFQAIFHMTYEFQLQEHFNPPRRYNPSSHVLSCACAGAVAAAT 180
Db 127 glgafysyrtqltmnlpfqshfityeflqeqvnprrtynpqsghisgglagaaaaat 186
QY 181 TPLDVCKTLNLTQESLALN-SHTGHTGMSAFRTYVQGVGTAYPRGVOARVIYQIPIS 239
Db 187 tpldvcktlntqenvalsnlsgrlsmanaftrtyvqglnglgyfkgigarviyqmps 246
QY 240 TAIAMSVYEFFKYLITKROEWRRA 263
Db 247 taiswsveffkylitkrqlenra 270
RESULT 6
ID AAM00938 standard; Protein; 366 AA.
XX AAM00938;
XX
XX 01-OCT-2001 (first entry)
XX
XX Human bone marrow protein, SEQ ID NO: 414.
XX
KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX
XX Homo sapiens.
XX
XX WO200153453-A2.
PN
XX 26-JUL-2001.
PD
XX 23-DEC-2000; 2000WO-US34960.
PF
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX WPI; 2001-488707/53.
DR N-PSDB; AAH90057.
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX Claim 10; Page 504-505; 648pp; English.
XX The present sequence is one of 251 novel human polypeptides encoded
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX Sequence 366 AA;
SQ
Query Match 71.1%; Score 999.5; DB 22; Length 366;
Best Local Similarity 68.6%; Pred. No. 3.1e-106;
Matches 181; Conservative 40; Mismatches 42; Indels 1; Gaps 1;
QY 1 MOSLPDPAARVNNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYELKKTLL 50
Db 101 mqslspdpkaqytslygaiklmrtegtwprlgrvnmimgsgpahamyfacyenmkrtl 160
QY 61 SDVIHPGNSHIANGAGCVATLLHDAAMNPAEVVKQRMQYNSPYHRTDCVRAVWONE 120
Db 161 ndvfhhqgnshlangjagsmatllhdavmnpaevvkqrlqmynsqhsaiscirtwrt 220
QY 121 GAGAFYRSYTTQLTWNVFPQAIHFMTYEFLOEHFNPQRYPNPSHVLSCACAGAVAAAT 180
Db 221 glgafyrtynpqtlmnpifqsthfityeflqeqvnphtynpqshisgglagalaat 280
QY 181 TPLDVCKTLNTQESALN-SHITGHITGMSAFRTYQVGVTA YFRGVQARVYQIPS 239
Db 281 tpldvcktlntqenvalsnisgrlsgmanaftrvqqlnglagfykgfgrarvlyqmps 340
QY 240 TAIAWSVYEFFKYLITTKRQEWRA 263
Db 341 taiswsyeffkylitkrqlenra 364
RESULT 7
AAU01989
ID AAU01989 standard; Protein; 174 AA.
XX AAU01989;
AC AAU01989;
XX 29-AUG-2001 (first entry)
XX Human secreted protein encoded by gene #25.
DE Human secreted protein; autoimmune disorder; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;
KW

KW nervous system disorder; bacterial infection; viral infection;
KW fungal infection; ocular disorder; wound healing; tissue regeneration;
KW epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.
XX Homo sapiens.
OS WO200123598-AL.
XX 05-APR-2001.
PD 26-SEP-2000; 2000WO-US26324.
PF 27-SEP-1999; 99US-0155807.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Komatsoulis G, Ruben SM, Rosen CA;
PI WPI; 2001-281684/29.
XX N-PSDB; AAS03906.
DR Forty one nucleic acid molecules encoding human secreted proteins, useful
DR in the prevention, treatment and diagnosis of cancer, immune disorders,
XX cardiovascular disorders and neurological diseases -
PT Disclosure; Page 501; 518pp; English.
XX The sequence represents a human secreted protein encoded by a nucleic
CC acid of the invention. Secreted proteins and their related nucleic acids
CC can be used in the diagnosis of or susceptibility to a pathological
CC condition by determining the presence or absence of a mutation in a
CC nucleic acid or the presence or amount of expression of a secreted
CC protein. The sequences are used to prevent, treat or ameliorate a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. The antibodies to the polypeptides can also be used in
CC alleviating symptoms associated with disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent
CC assays (ELISA). The disorders include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC fungi and ocular disorders e.g. corneal infection. The peptides can also
CC be used to aid wound healing and epithelial cell proliferation, to help
CC prevent skin ageing due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues, in chemotaxis and as a food
CC additive or preservative to alter storage capabilities.
XX Sequence 174 AA;
SQ
Query Match 65.1%; Score 916; DB 22; Length 174;
Best Local Similarity 100.0%; Pred. No. 4.3e-97;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 TLLHDAAMNPAEVVKQRMQYNSPYHRTDCVRAVWONEGAGAFYRSYTTQLTMTNVPQA 141
Db 1 tllhdaamnpaevvkqrmqynspyrhrtcdvrawwqnegagafysyttqmtmvpq 60
QY 142 IHFMTYEFLOEHFNPQRYPNPSHVLSCACAGAVAAATTPLDVCKTLNTQESALNSH 201
Db 61 ihfmyeflqehfnprypnpsshvlsqacagavaaattpldvcktlntqesalnsh 120
QY 202 ITGHITGMSAFRTYQVGVTA YFRGVQARVYQIPSTAIWSVYEFKYLIT 255
Db 121 itghitgmsaftrvyqvgvtayfrgvqarvlyqipstaiawsyeffkylit 174
RESULT 8
AAB40634
ID AAB40634 standard; Protein; 176 AA.
XX AAB40634;
AC AAB40634;

XX 08-FEB-2001 (first entry)
 XX Human ORFX ORF398 polypeptide sequence SEQ ID NO:796.
 DE
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antichyroid;
 KW antanaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 XX WO200058473-A2.
 PN
 XX
 PD 05-OCT-2000.
 XX
 XX 31-MAR-2000; 2000WO-US08621.
 PF
 XX 31-MAR-1999; 99US-0127607.
 PR
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 XX Shinkets RA, Leach W;
 PI WPI; 2000-602362/57.
 DR N-PSDB; AAC74843.
 DR
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 854; 5507pp; English.
 PS
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antichyroid; and antanaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 XX Sequence 176 AA;
 SQ

Query Match 64.7%; Score 910; DB 21; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2.2e-96;

Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 94 VVKORMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIHFMTYEFLOEH 153
 DB 5 VVKRMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIHFMTYEFLOEH 64
 QY 154 FNPQRRYNPSSHVLSGACAGAVAAATPLDVCKTLTNTQESLALNSHITGHTGMASAF 213
 DB 65 FNPQRRYNPSSHVLSGACAGAVAAATPLDVCKTLTNTQESLALNSHITGHTGMASAF 124
 QY 214 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 265
 DB 125 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 176
 RESULT 9
 AAB98221
 ID AAB98221 standard; Protein; 176 AA.
 XX
 AC AAB98221;
 XX
 XX 20-AUG-2001 (first entry)
 DT
 XX Human mitochondrial solute carrier (hMSC-homologue) protein SEQ:7.
 DE
 XX Human; hMSC; hMSC-homologue; mitochondrial solute carrier.
 KW
 XX Homo sapiens.
 OS
 XX CN1281895-A.
 PN
 XX 31-JAN-2001.
 PD
 XX 27-JUN-2000; 2000CN-0116795.
 PF
 XX 27-JUN-2000; 2000CN-0116795.
 PR
 XX (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
 PA
 XX Li N, Xu X, Xiao H;
 PI WPI; 2001-282781/30.
 DR N-PSDB; AAH22162.
 DR
 XX New human mitochondrial solute carrier (hMSC) protein isomer and its
 PT code sequence -
 PT
 XX Claim 4; Page 18 (disclosure); 22pp; Chinese.
 PS
 XX The present invention describes a human mitochondrial solute carrier
 CC protein, designated hMSC-homologue, which is expressed in human
 CC hypophysis. Also described are methods for the preparation and
 CC detection of hMSC-homologue protein and nucleotide sequences. The
 CC present sequence represents hMSC-homologue, as given in the present
 CC invention.
 XX
 XX Sequence 176 AA;
 SQ

Query Match 64.7%; Score 910; DB 22; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2.2e-96;
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 94 VVKORMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIHFMTYEFLOEH 153
 DB 5 VVKRMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIHFMTYEFLOEH 64
 QY 154 FNPQRRYNPSSHVLSGACAGAVAAATPLDVCKTLTNTQESLALNSHITGHTGMASAF 213
 DB 65 FNPQRRYNPSSHVLSGACAGAVAAATPLDVCKTLTNTQESLALNSHITGHTGMASAF 124
 QY 214 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 265
 DB 125 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 176
 RESULT 9
 AAB98221
 ID AAB98221 standard; Protein; 176 AA.
 XX
 AC AAB98221;
 XX
 XX 20-AUG-2001 (first entry)
 DT
 XX Human mitochondrial solute carrier (hMSC-homologue) protein SEQ:7.
 DE
 XX Human; hMSC; hMSC-homologue; mitochondrial solute carrier.
 KW
 XX Homo sapiens.
 OS
 XX CN1281895-A.
 PN
 XX 31-JAN-2001.
 PD
 XX 27-JUN-2000; 2000CN-0116795.
 PF
 XX 27-JUN-2000; 2000CN-0116795.
 PR
 XX (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
 PA
 XX Li N, Xu X, Xiao H;
 PI WPI; 2001-282781/30.
 DR N-PSDB; AAH22162.
 DR
 XX New human mitochondrial solute carrier (hMSC) protein isomer and its
 PT code sequence -
 PT
 XX Claim 4; Page 18 (disclosure); 22pp; Chinese.
 PS
 XX The present invention describes a human mitochondrial solute carrier
 CC protein, designated hMSC-homologue, which is expressed in human
 CC hypophysis. Also described are methods for the preparation and
 CC detection of hMSC-homologue protein and nucleotide sequences. The
 CC present sequence represents hMSC-homologue, as given in the present
 CC invention.
 XX
 XX Sequence 176 AA;
 SQ

Query Match 64.7%; Score 910; DB 22; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2.2e-96;
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 94 VVKORMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIHFMTYEFLOEH 153
 DB 5 VVKRMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIHFMTYEFLOEH 64
 QY 154 FNPQRRYNPSSHVLSGACAGAVAAATPLDVCKTLTNTQESLALNSHITGHTGMASAF 213
 DB 65 FNPQRRYNPSSHVLSGACAGAVAAATPLDVCKTLTNTQESLALNSHITGHTGMASAF 124
 QY 214 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 265
 DB 125 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 176
 RESULT 9
 AAB98221
 ID AAB98221 standard; Protein; 176 AA.
 XX
 AC AAB98221;
 XX
 XX 20-AUG-2001 (first entry)
 DT
 XX Human mitochondrial solute carrier (hMSC-homologue) protein SEQ:7.
 DE
 XX Human; hMSC; hMSC-homologue; mitochondrial solute carrier.
 KW
 XX Homo sapiens.
 OS
 XX CN1281895-A.
 PN
 XX 31-JAN-2001.
 PD
 XX 27-JUN-2000; 2000CN-0116795.
 PF
 XX 27-JUN-2000; 2000CN-0116795.
 PR
 XX (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
 PA
 XX Li N, Xu X, Xiao H;
 PI WPI; 2001-282781/30.
 DR N-PSDB; AAH22162.
 DR
 XX New human mitochondrial solute carrier (hMSC) protein isomer and its
 PT code sequence -
 PT
 XX Claim 4; Page 18 (disclosure); 22pp; Chinese.
 PS
 XX The present invention describes a human mitochondrial solute carrier
 CC protein, designated hMSC-homologue, which is expressed in human
 CC hypophysis. Also described are methods for the preparation and
 CC detection of hMSC-homologue protein and nucleotide sequences. The
 CC present sequence represents hMSC-homologue, as given in the present
 CC invention.
 XX
 XX Sequence 176 AA;
 SQ

Query Match 64.7%; Score 910; DB 21; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2.2e-96;

Db 125 rtvyvggtayfrvgarvlyqipstaiaawsvyeffkylithkrqeswragk 176

RESULT 10

ARB60658

ID AAB60658 standard; Protein; 331 AA.

XX

AC AAB60658;

XX

DT 04-MAY-2001 (first entry)

XX

DE Human mitochondrial solute carrier protein hMSC-o.

XX

KW Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;

KW preparation; detection.

XX

OS Homo sapiens.

XX

PN CN1269409-A.

XX

PD 11-OCT-2000.

XX

PF 17-MAR-2000; 2000CN-0114958.

XX

PR 17-MAR-2000; 2000CN-0114958.

XX

PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.

XX

PI Zhang X, Gao X, Xiao H;

XX

DR WPI; 2001-050544/07.

XX

DR N-PSDB; AAF59920.

XX

PT New human mitochondrion solute carrier protein and its nucleic acid -

XX

PS Claim 4; Page 20; 21pp; Chinese.

XX

CC The invention relates to a novel human mitochondrial carrier

CC protein, hMSC-o (AAB60658), and cDNA encoding it (AAF59920). hMSC-o is

CC expressed in normal human hypothalamus tissue. The invention also relates

CC to the preparation of hMSC-o proteins and nucleic acids, and the

CC detection of hMSC-o proteins and nucleic acids in a sample. The present

CC sequence represents hMSC-o.

XX

XX Sequence 331 AA;

XX

Query Match 60.3%; Score 848.5; DB 22; Length 331;

Best Local Similarity 68.4%; Pred. No. 6.7e-89;

Matches 154; Conservative 34; Mismatches 36; Indels 1; Gaps 1;

QY 1 MQSLQDPDPAARYRNVEALWRIITTEGLWRPRLGLNVTATGAGPAHALYFACYEKLKTL 60

Db 57 mqlslpdpkaqtsiygalckimrtgfwrlgrgvnmimgagpahamyfacyenmkrtl 116

QY 61 SDVHPGGNSHTANGAACGVATLLHDAAMNPAEVVKORMQYNPSVHRVTDVAVWONE 120

Db 117 ndvfhhqgushlanglagsmatllhdavmpaevvkqrlqmynsqhrsaiscirtvwrt 176

QY 121 GAGAFYRSYTTQLTMMNVFQAIHFMTYEFLOEHFNFPQRRYNPSSHVLGACAGAVAAAT 180

Db 177 glgafyrsyttqtmnpfqsibhfyeflqeqvnpbrtynpqshisgglagalaat 236

QY 181 TPLDVCKTLNLTQESLALN-SHITGHTCMASAFRTVYQGVTA 224

Db 237 tpldvcktlntqenvalslanislsgmanaftrvyqglnglpa 281

XX

AC AAU29748;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DT

XX

DE

XX

XX

KW

KW

XX

XX

OS

XX

PN

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

18-DEC-2001 (first entry)

Novel human secreted protein #239.

Human; vaccination; gene therapy; nutritional supplement;

stem cell proliferation; haematopoiesis; nerve tissue regeneration;

immune suppression; immune stimulation; anti-inflammatory; leukaemia.

Homo sapiens.

WO200179449-A2.

25-OCT-2001.

16-APR-2001; 2001WO-US08656.

18-APR-2000; 2000US-0552929.

26-JAN-2001; 2001US-0770160.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-611725/70.

Nucleic acids encoding a range of human polypeptides, useful in genetic

vaccination, testing and therapy -

Claim 20; Page 189; 765pp; English.

The invention relates to novel human secreted polypeptides. The

polypeptides and antibodies to the polypeptides are useful for

determining the presence of or predisposition to a disease associated

with altered levels of polypeptide. The polypeptides are also useful for

identifying agents (agonists and antagonists) that bind to them. Cells

expressing the proteins are useful for identifying a therapeutic agent

for use in treatment of a pathology related to aberrant expression or

physiological interactions of the polypeptide. Vectors comprising

the nucleic acids encoding the polypeptides and cells genetically

engineered to express them are also useful for producing the proteins.

The proteins are useful in genetic vaccination, testing and

therapy, and can be used as nutritional supplements. They may be used to

increase stem cell proliferation; to regulate haematopoiesis; and in

bone, cartilage, tendon and/or nerve tissue growth or regeneration;

immune suppression and/or stimulation; as anti-inflammatory agents; and

in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid

sequences of novel human secreted proteins of the invention.

Sequence 677 AA;

Query Match 53.7%; Score 754.5; DB 22; Length 677;

Best Local Similarity 66.7%; Pred. No. 1.3e-77;

Matches 138; Conservative 30; Mismatches 38; Indels 1; Gaps 1;

QY 1 MQSLQDPDPAARYRNVEALWRIITTEGLWRPRLGLNVTATGAGPAHALYFACYEKLKTL 60

Db 85 mqlslpdpkaqtsiygalckimrtgfwrlgrgvnmimgagpahamyfacyenmkrtl 144

QY 61 SDVHPGGNSHTANGAACGVATLLHDAAMNPAEVVKORMQYNPSVHRVTDVAVWONE 120

Db 145 ndvfhhqgushlanglagsmatllhdavmpaevvkqrlqmynsqhrsaiscirtvwrt 204

QY 121 GAGAFYRSYTTQLTMMNVFQAIHFMTYEFLOEHFNFPQRRYNPSSHVLGACAGAVAAAT 180

Db 205 glgafyrsyttqtmnpfqsibhfyeflqeqvnpbrtynpqshisgglagalaat 264

QY 181 TPLDVCKTLNLTQESLALN-SHITGHI 206

Db 265 tpldvcktlntqenvalslanislsgl 291


```
QY 140 QAIHMTYEFLEQHFNPORRYNPSSHVLSCAGACAGAAAAATPLDVCKTLTNTQESLALN 199
Db :||||:||||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 qsihityeflqevnphrtnypqshiisgglagalaattpldvcktlntqenvals 120
QY 200 -SHITGHTIGMASAFRTYVQGVGTAYFRGVOARVIYQIPSTAIWSVYEFKYLITRKQ 258
Db :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
121 lanisgrlsgmanaftrtyvqnglagyfkgiqarvyqmpstaiswsyeffkylitrq 180
QY 259 EEWRA 263
Db | ||
181 lenra 185

RESULT 14
AAB98222
ID AAB98222 standard; Protein; 187 AA.
XX
AC AAB98222;
XX
DT 20-AUG-2001 (first entry)
XX
DE Human mitochondrial solute carrier (hMSC) protein sequence.
XX
KW Human; hMSC; hMSC-homologue; mitochondrial solute carrier.
XX
OS Homo sapiens.
XX
PN CN1281895-A.
XX
PD 31-JAN-2001.
XX
PF 27-JUN-2000; 2000CN-0116795.
XX
PR 27-JUN-2000; 2000CN-0116795.
XX
PA (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
PI Li N, Xu X, Xiao H;
XX
WPI; 2001-282781/30.
XX
New human mitochondrial solute carrier (hMSC) protein isomer and its
code sequence -
XX
Example 2; Fig 3; 22pp; Chinese.
XX
The present invention describes a human mitochondrial solute carrier
protein, designated hMSC-homologue, which is expressed in human
hypophysis. Also described are methods for the preparation and
detection of hMSC-homologue protein and nucleotide sequences. The
present sequence represents the human mitochondrial solute carrier
hMSC protein sequence which is given in comparison with the
hMSC-homologue protein sequence in an example from the present
invention.
XX
SQ Sequence 187 AA;

Query Match 51.6%; Score 725.5; DB 22; Length 187;
Best Local Similarity 71.9%; Pred. No. 4.3e-75;
Matches 133; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY 80 VATLLHDAAMNPAEYVKQRMQYNSPYHRVTDVCRVAVQNEGAGAFYSYTTQLTMMNVPF 139
Db :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 matllhdavmpaevvkqrlgmynsqhrsaiscirtvrtgglgafysyttqltmnlpf 60

QY 140 QAIHMTYEFLEQHFNPORRYNPSSHVLSCAGACAGAAAAATPLDVCKTLTNTQESLALN 199
Db :||||:||||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 qsihityeflqevnphrtnypqshiisgglagalaattpldvcktlntqenvals 120
QY 200 -SHITGHTIGMASAFRTYVQGVGTAYFRGVOARVIYQIPSTAIWSVYEFKYLITRKQ 258
Db :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
121 lanisgrlsgmanaftrtyvqnglagyfkgiqarvyqmpstaiswsyeffkylitrq 180
```

```
QY 259 EEWRA 263
Db | ||
181 lenra 185

RESULT 15
AAB49665
ID AAB49665 standard; Protein; 187 AA.
XX
AC AAB49665;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human mitochondrial solute carrier (MSC) amino acid sequence.
XX
KW Mitochondrial solute carrier; MSC; human; adrenal gland.
XX
OS Homo sapiens.
XX
PN CN1268568-A.
XX
PD 04-OCT-2000.
XX
PF 02-MAR-2000; 2000CN-0111774.
XX
PR 02-MAR-2000; 2000CN-0111774.
XX
PA (SCHR-) SOUTH CHINA RES CENT NAT HUMAN GENE GROU.
XX
PI Li N, Qian B, Peng Y;
XX
WPI; 2001-050472/07.
XX
N-PSDB; AAC92501.
XX
New human mitochondrial solute carrier protein and its coding sequence
-
XX
Claim 2; Fig 2; 21pp; Chinese.
XX
The present invention relates to a human mitochondrial solute carrier
(MSC) protein represented by sequence AAB49665, which is encoded by cDNA
sequence AAC92501. The protein is expressed in adrenal gland tissue. The
invention includes methods for the preparation of the protein and
nucleic acid sequences.
XX
SQ Sequence 187 AA;

Query Match 51.6%; Score 725.5; DB 22; Length 187;
Best Local Similarity 71.9%; Pred. No. 4.3e-75;
Matches 133; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY 80 VATLLHDAAMNPAEYVKQRMQYNSPYHRVTDVCRVAVQNEGAGAFYSYTTQLTMMNVPF 139
Db :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 matllhdavmpaevvkqrlgmynsqhrsaiscirtvrtgglgafysyttqltmnlpf 60

QY 140 QAIHMTYEFLEQHFNPORRYNPSSHVLSCAGACAGAAAAATPLDVCKTLTNTQESLALN 199
Db :||||:||||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 qsihityeflqevnphrtnypqshiisgglagalaattpldvcktlntqenvals 120
QY 200 -SHITGHTIGMASAFRTYVQGVGTAYFRGVOARVIYQIPSTAIWSVYEFKYLITRKQ 258
Db :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
121 lanisgrlsgmanaftrtyvqnglagyfkgiqarvyqmpstaiswsyeffkylitrq 180
QY 259 EEWRA 263
Db | ||
181 lenra 185
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:58:50 ; Search time 48.11 Seconds
(without alignments)
134.541 Million cell updates/sec

Title: US-09-870-113-8

Perfect score: 1406

Sequence: 1 MQSLQDPDPAARNVLEALW.....VYEFKYLITKROENRAGK 265

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	429.5	30.5	289	4	US-09-068-140A-10
2	429.5	30.5	328	4	US-09-068-140A-15
3	251.5	17.9	351	2	US-08-933-750C-19
4	251.5	17.9	351	4	US-09-234-613-19
5	219.5	15.6	469	4	US-09-188-930-339
6	213	15.1	447	4	US-09-160-119-4
7	213	15.1	674	4	US-09-160-119-2
8	212.5	15.1	320	2	US-08-933-750C-12
9	212.5	15.1	320	4	US-09-234-613-12
10	168.5	12.0	432	2	US-08-937-466-4
11	168.5	12.0	432	2	US-09-172-528-4
12	168.5	12.0	432	3	US-09-318-199-4
13	168.5	12.0	432	4	US-09-503-579-4
14	168	11.8	311	2	US-08-775-009-32
15	166.5	11.8	311	2	US-08-775-009-33
16	164.5	11.7	308	2	US-08-937-466-2
17	164.5	11.7	308	2	US-09-172-528-2
18	164.5	11.7	308	3	US-09-318-199-2
19	164.5	11.7	308	4	US-09-503-579-2
20	163.5	11.6	306	5	PCT-US94-09799-1
21	161	11.5	312	2	US-09-142-565-2
22	154	11.0	307	2	US-08-807-861A-56
23	154	11.0	307	3	US-09-210-681-56
24	154	11.0	307	3	US-08-946-719A-56
25	148	10.5	256	2	US-08-937-466-6
26	148	10.5	256	2	US-09-172-528-6
27	148	10.5	256	3	US-09-318-199-6

28 148 10.5 256 4 US-09-503-579-6 Sequence 6, Appli
29 137 9.7 299 1 US-08-518-878B-56 Sequence 56, Appl
30 137 9.7 299 2 US-08-470-868A-56 Sequence 56, Appl
31 137 9.7 309 1 US-08-518-878B-51 Sequence 51, Appl
32 137 9.7 309 2 US-08-807-861A-51 Sequence 51, Appl
33 137 9.7 309 2 US-08-470-868A-51 Sequence 51, Appl
34 137 9.7 309 3 US-09-210-681-51 Sequence 51, Appl
35 137 9.7 309 3 US-08-946-719A-51 Sequence 51, Appl
36 121.5 8.6 298 3 US-08-961-871-10 Sequence 10, Appl
37 121.5 8.6 312 4 US-09-188-930-142 Sequence 142, Appl
38 106 7.5 303 1 US-08-518-878B-37 Sequence 37, Appl
39 106 7.5 303 1 US-08-294-522B-36 Sequence 36, Appl
40 106 7.5 303 2 US-08-807-861A-37 Sequence 37, Appl
41 106 7.5 303 2 US-08-470-868A-37 Sequence 37, Appl
42 106 7.5 303 3 US-09-210-681-37 Sequence 37, Appl
43 106 7.5 303 3 US-08-946-719A-37 Sequence 37, Appl
44 85 6.0 1832 3 US-09-335-409-4 Sequence 4, Appli
45 85 6.0 1832 4 US-09-568-102-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-068-140A-10
; Sequence 10, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409 September 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinher, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-10

```
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-15

Query Match 30.5%; Score 429.5; DB 4; Length 289;
Best Local Similarity 38.7%; Pred. No. 1.6e-42;
Matches 99; Conservative 38; Mismatches 108; Indels 11; Gaps 4;

QY 9 AARYRNVLALWRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVIHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 SAQSAGLRQALGASILKVEGPAGLYRGIGAMGLGAGPAHAHVFSVYEMCKETFS---HGDP 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 NSHIANGAAGCVATLLHDAAMNPAEYVVKORMQYNSPYHRVTDVCRVAVQNEGAGAFYRS 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 SNGAHAVSGVFATVADAVITPMDVVKQRLQSSPYKGVVDVCRVRLVEEGIGAFYAS 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 YTTQLTMNVPFOAIHPTWTFEFLQE---HFNPO---RRYNPSSHVLSGACAGAVAAAATPPL 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 YRTTVVMNAPFTAVHPTATYEATKKGLEVPETANDENLLVHATAGAAALAAVVTTPPL 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 DVCKTLLNTQESLALNSHTGHTGMASAFRTVYQVGGVTATYFRGVOARVYIOPSTATA 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 DVVKTOLQCGVCGCDRFSSSSIQDVIG---SIVKNGYVGLMRGWIPRMLFHAPAAAI 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 WSVYEFFKYLITKROE 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 WSTYEASKTFFOKLNE 281
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: NO. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-15

Query Match 30.5%; Score 429.5; DB 4; Length 328;
Best Local Similarity 38.7%; Pred. No. 2e-42;
Matches 99; Conservative 38; Mismatches 108; Indels 11; Gaps 4;

QY 9 AARYRNVLALWRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVIHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 SAQSAGLRQALGASILKVEGPAGLYRGIGAMGLGAGPAHAHVFSVYEMCKETFS---HGDP 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 NSHIANGAAGCVATLLHDAAMNPAEYVVKORMQYNSPYHRVTDVCRVAVQNEGAGAFYRS 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 SNGAHAVSGVFATVADAVITPMDVVKQRLQSSPYKGVVDVCRVRLVEEGIGAFYAS 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 YTTQLTMNVPFOAIHPTWTFEFLQE---HFNPO---RRYNPSSHVLSGACAGAVAAAATPPL 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 YRTTVVMNAPFTAVHPTATYEATKKGLEVPETANDENLLVHATAGAAALAAVVTTPPL 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 DVCKTLLNTQESLALNSHTGHTGMASAFRTVYQVGGVTATYFRGVOARVYIOPSTATA 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 DVVKTOLQCGVCGCDRFSSSSIQDVIG---SIVKNGYVGLMRGWIPRMLFHAPAAAI 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 WSVYEFFKYLITKROE 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 WSTYEASKTFFOKLNE 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-08-933-750C-19
; Sequence 19, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536.
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
```


Db 218 ASRSNNNCIVGGFTQMTIREGAKSLWRG-NGINVLK--IAPESAIKFMAYEQMKRLVGS 274
QY 63 V-----IHPGNSHTANGAAGCVATLLHDAAMNPAEVVKQRMOMYNS-PYHRVTDVCRVAV 117
Db 275 QETRIH---ERLVAGSLAGAIA-----OSSYPMEVLUKTRMALKTKQYSGMLDCARRIL 327
QY 118 QNEGAGAFYRSYTTQLTMTNVPFOAIHFMTEYEFLOEHPNORRY-----NPSSHVL--SGA 170
Db 328 AKEGVAAFYKGYIPNMLGIIPYAGIDLAVYETLKNTW--LQRYAVNSADPGVFLACGT 385
QY 171 CAGVAAAATTPLDVCKTLNTQESLALNSHITGTMASAFRTYVQGVGTAYFRGVQ 230
Db 386 ISSTCGQLASYPLALVTRMQAASIEGAPEVT-----MSSLFKQILRTGAGFLYRGLA 440
QY 231 ARVIYQIPSTAIAMSVYEFK 251
Db 441 PFMKVIPAVSISVYVENLK 461

RESULT 6
US-09-160-119-4
; Sequence 4, Application US/09160119A
; Patent No. 6316219
; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRIL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160,119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 447
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-4

Query Match 15.1%; Score 213; DB 4; Length 447;
Best Local Similarity 26.7%; Pred. No. 1.4e-16;
Matches 68; Conservative 41; Mismatches 110; Indels 36; Gaps 8;
QY 12 YRNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLS DVI-----HP 66
Db 143 YKNSFDCFKKVLRYEGFFGLYRGLLPOLLGVAPEKA-----IKLTVNDVDRKFMHK 194
QY 67 GGN-----SHIANGAAGCVATLLHDAAMNPAEVVKQRMOMYNSPYHRVTDVCR-----AVW 117
Db 195 DGSVPLAAEILAGCAGGSQVIF-----TNPLEIVKIRLQVAG-----EITGPRVSALS 246
QY 118 QNEGAGAFYRSYTTQLTMTNVPFOAIHFMTEYEFLOEHP-NPQRRYNPSSHVLSGACAGAVA 176
Db 247 RDLGFFGIYKGACACFLRDIPFSALYPCYAHVKASFANEDGVSPGSLLAGAIGMPA 306
QY 177 AAATTPLDVCKTLNTQESLALNSHITGTMASAFRTYVQGVGTAYFRGVQARVIYQ 236
Db 307 ASLVTADVIRKTRLQVAARAGQTTY-----SGVIDCFKILREEGPKALKWGAGARVFRS 361
QY 237 IPSTAIAMSVYEFK 251
Db 362 SPOFGVTLTYELLQ 376

RESULT 7
US-09-160-119-2
; Sequence 2, Application US/09160119A
; Patent No. 6316219

; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRIL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160,119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 674
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-2

Query Match 15.1%; Score 213; DB 4; Length 674;
Best Local Similarity 26.7%; Pred. No. 2.7e-16;
Matches 68; Conservative 41; Mismatches 110; Indels 36; Gaps 8;
QY 12 YRNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLS DVI-----HP 66
Db 370 YKNSFDCFKKVLRYEGFFGLYRGLLPOLLGVAPEKA-----IKLTVNDVDRKFMHK 421
QY 67 GGN-----SHIANGAAGCVATLLHDAAMNPAEVVKQRMOMYNSPYHRVTDVCR-----AVW 117
Db 422 DGSVPLAAEILAGCAGGSQVIF-----TNPLEIVKIRLQVAG-----EITGPRVSALS 473
QY 118 QNEGAGAFYRSYTTQLTMTNVPFOAIHFMTEYEFLOEHP-NPQRRYNPSSHVLSGACAGAVA 176
Db 474 RDLGFFGIYKGACACFLRDIPFSALYPCYAHVKASFANEDGVSPGSLLAGAIGMPA 533
QY 177 AAATTPLDVCKTLNTQESLALNSHITGTMASAFRTYVQGVGTAYFRGVQARVIYQ 236
Db 534 ASLVTADVIRKTRLQVAARAGQTTY-----SGVIDCFKILREEGPKALKWGAGARVFRS 588
QY 237 IPSTAIAMSVYEFK 251
Db 589 SPOFGVTLTYELLQ 603

RESULT 8
US-08-933-750C-12
; Sequence 12, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0


```
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-937-466-4
```

```
Query Match 12.0%; Score 168.5; DB 2; Length 432;
Best Local Similarity 24.6%; Pred. No. 2.5e-11;
Matches 66; Conservative 38; Mismatches 121; Indels 43; Gaps 10;

Qy 9 AARYNVLEALWRIIRTEGLRPMRLGNVTATGAGPAHALYFA-----CYEKLKKTLSDV 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 SVQYRGVLGTILTMVTEGRSPSYSGL-----VAGLHRQMSFASIRIGLYDSVKQFYT-- 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 64 IHPGNSHIANG---AAGCVATLLHDAAMNPAEVVKQRMQY-----NSPYHRVTDVCV 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 --PKGADHSSVAIRLAGCTTGAMAVTCAQPTDVVVKVRFQAMIRLGTGGERKYRGTMDAY 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 114 RAVMNEGAGAFYRSYTTQLTMMNVFPQAIHFMTYEFLEQHEFNPPORRY--NPSSHVLSCAC 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 RTIAREEGVRLGWKTWPNITRNAIVNCAEMVTYDIIEKELLESHLFTDNFCHFVSAGF 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 172 AGAVAAATATPLDVCCKTLNTQESLALNSHITGHTMASAFRTVYQGGVTAYFRGVQA 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 AGFCATVVASPDVVVKTRY-----MNAPLGRYSPLHCLMKWVAQ-EGPTAFYKGF-- 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 232 RVYQIPS--TAIAWSVYEFKYLITKR 257
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 -----VPSFLRLGAWNMVMFVTEOLKR 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 11
US-09-172-528-4
; Sequence 4, Application US/09172528
; Patent No. 5952469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/172,528
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-172-528-4

Query Match 12.0%; Score 168.5; DB 2; Length 432;
Best Local Similarity 24.6%; Pred. No. 2.5e-11;
Matches 66; Conservative 38; Mismatches 121; Indels 43; Gaps 10;

Qy 9 AARYNVLEALWRIIRTEGLRPMRLGNVTATGAGPAHALYFA-----CYEKLKKTLSDV 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 SVQYRGVLGTILTMVTEGRSPSYSGL-----VAGLHRQMSFASIRIGLYDSVKQFYT-- 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 64 IHPGNSHIANG---AAGCVATLLHDAAMNPAEVVKQRMQY-----NSPYHRVTDVCV 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 --PKGADHSSVAIRLAGCTTGAMAVTCAQPTDVVVKVRFQAMIRLGTGGERKYRGTMDAY 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 114 RAVMNEGAGAFYRSYTTQLTMMNVFPQAIHFMTYEFLEQHEFNPPORRY--NPSSHVLSCAC 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 RTIAREEGVRLGWKTWPNITRNAIVNCAEMVTYDIIEKELLESHLFTDNFCHFVSAGF 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 172 AGAVAAATATPLDVCCKTLNTQESLALNSHITGHTMASAFRTVYQGGVTAYFRGVQA 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 AGFCATVVASPDVVVKTRY-----MNAPLGRYSPLHCLMKWVAQ-EGPTAFYKGF-- 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 232 RVYQIPS--TAIAWSVYEFKYLITKR 257
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 -----VPSFLRLGAWNMVMFVTEOLKR 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 12
US-09-318-199-4
; Sequence 4, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```


APPLICATION NUMBER: US/08/775,009
FILING DATE: 27-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CH-0681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-009-32

Query Match 11.9%; Score 168; DB 2; Length 311;
Best Local Similarity 24.2%; Pred. No. 1.7e-11;
Matches 65; Conservative 38; Mismatches 118; Indels 48; Gaps 8;
QY 11 RYRVLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTLSDVIHPGNS 70
DB 63 RYRGIGDCVRQTVRSHGLYRGLSLLYGSIPKAAVRFGMFEFLSNHMDA---QGR 119
QY 71 HIANG-----AAGCVATLLHDAAMNPAEVVKQRM-----QMYNSPYHRTDCVR 117
DB 120 DSTRGLLGLGAGVAAV---VCPMTVKVKFIHDTSPNPKYRGFFHGVREIYR--- 173
QY 118 QNEGAGAFYSRTTQLTMMNVPFOAIHFMTYEFLEQHF---NFORRYNPSSHVLGACAGA 174
DB 174 -EGLKGTGYOGLTATVLKOGSNOAIRFFVMTSLRNWYRGDNPNKPNPLITGVFGAIGA 232
QY 175 VAAATTPLDVCKTLNTOESLALNSHITGTMASAFRTYVQVG-----GVTA 227
DB 233 ASVFGNTPLDVAKTRMOGLE-----AHKYRNTWDCGLQILKKEGLKAFYK 277
QY 228 GVOARVIYQIPSTAIWSVYEFKYLITK 256
DB 278 GTFPRGRVCLDVAIVFIYDEVYKLLNK 306

RESULT 15
US-08-775-009-33
Sequence 33, Application US/08775009
Patent No. 5935783
GENERAL INFORMATION:
APPLICANT: Gong, Wellong
APPLICANT: Emanuel, Beverly S.
APPLICANT: Budarf, Marcia L.
APPLICANT: Roe, Bruce
TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
TITLE OF INVENTION: Velociocardiofacial Syndrome Minimal Critical Region
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5935783ris, LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,009
FILING DATE: 27-DEC-1996

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CH-0681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-009-33
Query Match 11.8%; Score 166.5; DB 2; Length 311;
Best Local Similarity 24.7%; Pred. No. 2.6e-11;
Matches 69; Conservative 35; Mismatches 124; Indels 51; Gaps 8;
QY 4 LQDPDPA---RYRVLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTL 60
DB 53 LQIDERANPPRYRGIGDCVRQTVRSHGLYRGLSLLYGSIPKAAVRFGMFEFLSNHM 112
QY 61 SDV-----IHPGNSHIANGAAGCVATLLHDAAMNPAEVVKQRM-----QMYNSPYH 107
DB 113 RDAQGLDSRGLLGLGAGVAAV---PMTVKVKFIHDTSSNPKNKRGFFH 166
QY 108 RYTDVAVYQNEGAGAFYSRTTQLTMMNVPFOAIHFMTYEFLEQHF---NFORRYNPSS 164
DB 167 GVREIYR---EGLKGTGYOGLTATVLKOGSNOAIRFFVMTSLRNWYQDNPNKPNPLI 222
QY 165 HVLGACAGAVAAATTPLDVCKTLNTOESLALNSHITGTMASAFRTYVQVG----- 220
DB 223 TGVEGAVAGAASVFGNTPLDVAKTRMOGLE-----AHKYRNTLDCGVQL 267
QY 221 ---GVTAVERGVQARVIYQIPSTAIWSVYEFKYLITK 256
DB 268 KNEGPKAFYKGVPRGRVCLDVAIVFIYDEVYKLLNK 306

Search completed: August 27, 2002, 03:58:51
Job time: 4893 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:57:41 ; Search time 80.04 Seconds
(without alignments)
318.137 Million cell updates/sec

Title: US-09-870-113-8
Perfect score: 1406
Sequence: 1 MQSLQDPDAARYNVLEALW.....VVEFFKYLITKQEEWRAGK 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	677	48.2	312	2 T26089	hypothetical prote
2	467	33.2	303	2 T39149	probable RNA splic
3	441	31.4	304	2 S13533	mtRNA splice defec
4	431.5	30.7	781	2 A86205	hypothetical prote
5	428	30.4	314	2 S55179	mtRNA splice defec
6	414	29.4	331	2 T00582	probable mitochond
7	283	20.1	328	2 T19322	hypothetical prote
8	275.5	19.6	368	2 S54524	hypothetical prote
9	251.5	17.9	351	2 T43493	hypothetical prote
10	248.5	17.7	367	2 G96770	hypothetical prote
11	242.5	17.2	384	1 S44092	probable carrier p
12	226	16.1	284	2 S45458	pet8 protein - yea
13	224.5	16.0	902	1 S54495	probable carrier p
14	219.5	15.6	269	2 T20290	hypothetical prote
15	216.5	15.4	330	2 T09362	hypothetical prote
16	213	15.1	336	2 T01839	hypothetical prote
17	210	14.9	352	2 T01729	mitochondrial solu
18	209.5	14.9	436	2 J01459	Btl protein precur
19	209.5	14.9	475	2 T50886	peroxisomal Ca-dep
20	208.5	14.8	392	2 T05350	adenylate transloc
21	208	14.8	366	2 S64589	hypothetical prote
22	206	14.7	322	2 S57116	probable carrier p
23	204	14.5	271	2 S62485	probable mitochond
24	203.5	14.5	338	2 T40968	probable mitochond
25	201.5	14.3	349	2 B96753	hypothetical prote
26	200	14.2	479	2 T49871	peroxisomal Ca-dep
27	199.5	14.2	307	2 S60949	probable phosphate
28	198	14.1	345	2 T37576	probable mitochond
29	196.5	14.0	326	2 S57544	probable membrane

30 196 13.9 358 2 D84901
31 193 13.7 358 2 T45934
32 192 13.7 335 2 T50393
33 191 13.6 588 2 T22688
34 188 13.4 288 2 T51595
35 188 13.4 309 2 T48156
36 187.5 13.3 296 2 B96830
37 186 13.2 332 2 T47703
38 186 13.2 373 2 S48451
39 185 13.2 325 2 T04273
40 184 13.1 300 2 S50556
41 184 13.1 335 2 S50453
42 181 12.9 500 2 T39385
43 179.5 12.8 650 2 T32897
44 179 12.7 258 2 T39347
45 178 12.7 277 2 T40033

ALIGNMENTS

RESULT 1

T26089

hypothetical protein W02B12.9 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C;Accession: T26089

R;Swinburne, J.; Ahnscough, R.

submitted to the EMBL Data Library, October 1995

A;Reference number: Z20149

A;Accession: T26089

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-312 <WIL>

A;Cross-references: EMBL:Z66521; PIDN:CAA91399.1; GSPDB:GN00020; CESP:W02B12.9

A;Experimental source: clone W02B12

C;Genetics:

A;Gene: CESP:W02B12.9

A;Map position: 2

A;Introns: 18/3; 251/3; 286/3

C;Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 48.2%; Score 677; DB 2; Length 312;

Best Local Similarity 50.6%; Pred. No. 7.7e-54;

Matches 133; Conservative 47; Mismatches 73; Indels 10; Gaps 3;

QY 1 MQSLQDPDAARYNVLEALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKKKTL 60

Db 45 MQSLCPCEPKCTPPVHSLMSIVKREGWLRPLRGVNAVAGSMFAHALYFTVYEKMGYL 104

QY 61 SDVIHPGNGSH-IANGAAGCVATLLHDAAMNPAEVVKORMQYNSPYHRVTDVCYRAYWQN 119

Db 105 TG--NSAGHSNTLAYGASGVVATLIHDAIMNPAEVVKORMQAFSPYSSLSGCARCVNR 162

QY 120 EGAGAFYRSYTTQITMNVPPQAHFMYEYFLEQHFNPORRYNPSSHVLSCACAGAVAAA 179

Db 163 EGVAAFYRSYTTQIAMNVPPQAHFMYEYFLEQHFNPORRYNPSSHVLSCACAGAVAAA 222

QY 180 TPLDVCCKTLINTQESLANS-----HITGHITGMASAFRTVYQVGGVTAYEFGVOAR 232

Db 223 TTPMDCVKTVLNTQQAADPANRIFLQARYRYRGISDAVRTIYSORGLSGFSCGLQAR 282

QY 233 VVIQIPSTAIWSYEFKYLIT 255

Db 283 VIFQVPATALSWSYELFKFMLS 305

RESULT 2

T39149

probable RNA splicing proteinmitochondial carrier protein - fission yeast (Schizosacc

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: T39149
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21748
A:Accession: T39149
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-303 <OLI>
A:Cross-references: EMBL:Z99168; PIDN:CAB16300.1; GSPDB:GN00066; SPDB:SPAC8C9.12c
A:Experimental source: strain 972h-; cosmid c8c9
C:Genetics:
A:Gene: SPDB:SPAC8C9.12c
A:Map position: 1
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 33.2%; Score 467; DB 2; Length 303;
Best Local Similarity 37.0%; Pred. No. 8.9e-35;
Matches 90; Conservative 55; Mismatches 92; Indels 6; Gaps 2;
QY 14 NVLEALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKLTLSVDVHPGGNSHTA 73
DB 60 NIVNSVIKISSIEGVYSLWRGSISSVMGAPSHAIYFSVLEFFSK-----INASPDRLPA 115
QY 74 NGAAGCVATLLHDAAMNPAEVVKQRMQMYNSPYHRVTDCAVRAVQNEGAGAFYRSYTTQL 133
DB 116 SALAGACAITISDAFMTDPDVIKQRMQLPSRKYKSAHCAITTVFERNELGAFYISYPTCI 175
QY 134 TNWVPQAIHFMTYEFELQEHFNQRRYNPSSHVLSGACAGAAVAAATPLDVCCKILLNQ 193
DB 176 AMSIPTAIQVATYDTCMSFLNPNAVDPYDTSIISSGLSGAIASSLTPLDVLVVKLLQTR 235
QY 194 ESLANSHITGHTGMASAFRTVYQVGGVTAYFRGVQARVIYQIPSTALAWSVVEFFKYL 253
DB 236 GSSSTPE--VRKCKSLDVRVIYNGGIPSPFKGIRPRWVAMPATAVSWAAYEAGKEI 293
QY 254 ITK 256
DB 294 LIR 296

RESULT 3
S13533
mRNA splice defect-suppressing mitochondrial carrier MRS4 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKR052c
C:Species: Saccharomyces cerevisiae
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jul-2000
C:Accession: S13533; S38126
R:Wiesenberger, G.; Link, T.A.; von Ahsen, U.; Waldbherr, M.; Schweyen, R.J.
J. Mol. Biol. 217, 23-37, 1991
A:Title: MRS3 and MRS4, two suppressors of mRNA splicing defects in yeast, are new members of the ADP/ATP carrier family
A:Reference number: S13532; MUID:91108815
A:Accession: S13533
A:Molecule type: DNA
A:Residues: 1-304 <JMO>
A:Cross-references: EMBL:X56444; NID:g3995; PIDN:CAA39828.1; PID:g3996
R:Visiers, S.; Urrestarazu, L.A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38118
A:Accession: S38126
A:Molecule type: DNA
A:Residues: 1-304 <VIS>
A:Cross-references: EMBL:Z28277; NID:g486506; PID:g486507; MIPS:YKR052c
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:MRS4
A:Cross-references: SGD:S0001760; MIPS:YKR052c
A:Map position: 11R

A:Genome: nuclear
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: duplication; mitochondrion; transmembrane protein
F:20-109/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:117-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>

F:206-301/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 31.4%; Score 441; DB 2; Length 304;
Best Local Similarity 38.4%; Pred. No. 2.1e-32;
Matches 93; Conservative 41; Mismatches 106; Indels 2; Gaps 1;

QY 15 VLEALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKLTLSVDVHPGGNSHIAN 74
DB 64 MISQISKISTMEGSMALMGVOSVILGAGPAHAVFGTYEFCFKARLISPEDQTHQPMKT 123
QY 75 GAAGCVATILLHDAAMNPAEVVKQRMQMYNSPYHRVTDCAVRAVQNEGAGAFYRSYTTQLT 134
DB 124 ALSGTIATIAADALNPEDTVKQRLQDNTN--LRVWNVTKQIYQNEGFAAFYSYPTTLA 181
QY 135 MNVPQAIHFMTYEFELQEHFNQRRYNPSSHVLSGACAGAAVAAATPLDVCCKILLNQ 194
DB 182 MNIPFAAFNFMITYESAKFFNPQNSYNPLIHCLCGGISGATCAALTPLDCKITVLQVRG 241
QY 195 SLALNSHITGHTGMASAFRTVYQVGGVTAYFRGVQARVIYQIPSTALAWSVVEFFKYL 254
DB 242 SETVSEIEMKDANTGRASRAILLEVHGKGFWRGLKPRIVANIPATAISWTAYEAKHFL 301
QY 255 TK 256
DB 302 MK 303

RESULT 4
A86205
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86205
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: A86205
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-781 <STO>
A:Cross-references: GB:AB005172; NID:g8954043; PIDN:AAF82217.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 30.7%; Score 431.5; DB 2; Length 781;
Best Local Similarity 40.0%; Pred. No. 5e-31;
Matches 104; Conservative 35; Mismatches 104; Indels 17; Gaps 6;

QY 1 MQSLOPDAARYNVLEALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKLT 60
DB 64 MQALRCP- LKPVGIREAFRSIIQKEGSPSALYRGIAWMLGAGPAHAHYVFSFYSKYL 122
QY 61 SDVIHPGG--NSHIANGAGCVATILLHDAAMNPAEVVKQRMQMYNSPYHRVTDCAVRAVWO 118
DB 123 S-----AGQNNVAHMSGVGFATISSDAVFTPDNDVQKRLQMGEGTYKGVWDCKVRVLR 177
QY 119 NEGAGAFYRSYTTQLTMVVPFOAHFMTYEFLOE---HFNQRRYNPSS---HVLSGACA 172
DB 178 EEGIGAFYASRYTTVMNAPTAVHFATYEAARKGLMEFSPDRISDESGWLVAHATAGAA 237
QY 173 GAVAAAAATPLDVCCKILLNTQESIALNSHITGHTGMASAFRTVYQVGVGTAYFRGVQAR 232
DB 238 GGLAAAVTPLDVLVVKTLQCCQVCVCGRFTSSSI---SHVLRITVKKDGYRGLLRGLR 294

Query Match 15.4%; Score 216.5; DB 2; Length 330;
Best Local Similarity 33.7%; Pred. No. 5.7e-12;
Matches 70; Conservative 21; Mismatches 94; Indels 23; Gaps 7;
QY 26 EGLWRPMRGLNVATGAGPAHALYFACYE---KLKTLSDVTHPPGNSHANGACGVA 81

Db 94 KGLY---SGLAGNIAGVLPASALFVGVEPTKOKLLKTPPD--HLSAVAHLTAGAIGGLA 148
QY 82 TLLHDAAMNPAEVLKORMQMYNSPYHRVTDVRAVQWNECAGAFYRSYTTQLTMTNVPFQA 141
Db 149 ASL---IRVPTVVVKORMQ--TGOFTSAPSAVRMIASKEGFRGLYAGYRSFLLRDLFPDA 203
QY 142 IHFMTYEFLQEHFN--PQRRYNPSSHVLSGACAGAVAAAATPLDVCKTLLNTQESLALN 199
Db 204 IQFCIYEQLCLGYKKAARRELSDPENALICAFAGALTGAVTTPLDVKTFLMWQGS---- 259
QY 200 SHITGHITGMASAPRTYVQVGGVTAYFR 227
Db 260 ---AKOYQIVDCVQTVIREEGAPALLK 284

Search completed: August 27, 2002, 03:57:42
Job time: 4948 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:16:46 ; Search time 37.5 Seconds
(without alignments)
273.618 Million cell updates/sec

Title: US-09-870-113-8

Perfect score: 1406

Sequence: 1 MQSLQDPDAARYRNLEALW.....VYEFFKYLITKROEWRAGK 265

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	441	31.4	304	1 MRS4_YEAST	P23500 saccharomyc
2	428	30.4	314	1 MRS3_YEAST	P10566 saccharomyc
3	283	20.1	328	1 Y051_CAEEL	Q09461 caenorhabdi
4	275.5	19.6	368	1 YK39_YEAST	Q03829 saccharomyc
5	251.5	17.9	359	1 CG69_HUMAN	Q9BZJ4 homo sapien
6	226	16.1	284	1 PFB2_YEAST	P38921 saccharomyc
7	217	15.4	676	1 CMC2_MOUSE	Q9GXX4 mus musculu
8	213	15.1	675	1 CMC2_HUMAN	Q9HJ21 homo sapien
9	212.5	15.1	320	1 DNC_HUMAN	Q9HC21 homo sapien
10	210	14.9	695	1 CMC1_DROME	Q9VA73 drosophila
11	209.5	14.9	436	1 B71_MAIZE	P29518 zea mays (m
12	208	14.8	366	1 YG5F_YEAST	P53320 saccharomyc
13	206	14.7	322	1 SFC1_YEAST	P33303 saccharomyc
14	204	14.5	271	1 YAD8_SCHPO	Q09834 schizosacch
15	199.5	14.2	307	1 ODC2_YEAST	Q99297 saccharomyc
16	198	14.1	345	1 YDE9_SCHPO	Q10442 schizosacch
17	196.5	14.0	315	1 MFT_HUMAN	Q9H2D1 homo sapien
18	191	13.6	588	1 CMC2_CAEEL	Q20799 caenorhabdi
19	189.5	13.5	301	1 ORT1_MOUSE	Q9WVD5 mus musculu
20	187	13.3	678	1 CMC1_HUMAN	O75746 homo sapien
21	186	13.2	315	1 SA18_HUMAN	Q9H1K4 homo sapien
22	186	13.2	373	1 YTA6_YEAST	P40556 saccharomyc
23	184	13.1	300	1 YEO3_YEAST	P40035 saccharomyc
24	184	13.1	335	1 YEA6_YEAST	P39953 saccharomyc
25	183.5	13.1	301	1 MCAT_HUMAN	Q43772 homo sapien
26	183	13.0	312	1 SA18_MOUSE	Q9DB41 mus musculu
27	176	12.5	301	1 MCAT_RAT	P97521 rattus norv
28	174.5	12.4	301	1 ORT2_HUMAN	Q9BXI2 homo sapien
29	174.5	12.4	324	1 PMT_YEAST	P32332 saccharomyc
30	173	12.3	330	1 GDC_BOVIN	Q01888 bos taurus
31	172	12.2	587	1 CMC3_CAEEL	Q19529 caenorhabdi
32	170.5	12.1	311	1 UCP3_CANFA	Q9N2I9 canis famil
33	168.5	12.0	302	1 YD1K_SCHPO	Q10248 schizosacch

34	168.5	12.0	308	1 UCP3_MOUSE	P56501 mus musculu
35	168	11.9	311	1 TXTP_HUMAN	P53007 homo sapien
36	167	11.9	332	1 GDC_HUMAN	P16260 homo sapien
37	166.5	11.8	311	1 TXTP_RAT	P32089 rattus norv
38	165.5	11.8	311	1 TXTP_BOVIN	P79110 bos taurus
39	165	11.7	312	1 TXTP_CAEEL	P34519 caenorhabdi
40	164.5	11.7	301	1 ORT1_HUMAN	Q9Y619 homo sapien
41	164.5	11.7	308	1 UCP3_PIG	O97649 sus scrofa
42	164.5	11.7	308	1 UCP3_RAT	P56499 rattus norv
43	163	11.6	292	1 ORT1_YEAST	Q12375 saccharomyc
44	163	11.6	702	1 CMC1_CAEEL	Q21153 caenorhabdi
45	162.5	11.6	311	1 MPCP_YEAST	P23641 saccharomyc

ALIGNMENTS

```

RESULT 1
MRS4_YEAST
ID MRS4_YEAST STANDARD; PRT; 304 AA.
AC P23500;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Mitochondrial RNA splicing protein MSR4.
GN MSR4 OR YKR052C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M1301.
RX MEDLINE=91108615; PubMed=1703236;
RA Wiesenberger G., Link T.A., von Ahlsen U., Waldherr M., Schweyen R.J.;
RT "MRS3 and MRS4, two suppressors of mRNA splicing defects in yeast,
RL J. Mol. Biol. 217:23-37(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Visser S., Urrestazu L.A., Jauniaux J.-C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MRS4 SUPPRESSES A MITOCHONDRIAL SPLICE DEFECT IN THE
CC FIRST INTRON OF THE COB GENE. IT MAY ACT AS A CARRIER, EXERTING
CC ITS SUPPRESSOR ACTIVITY VIA MODULATION OF SOLUTE CONCENTRATIONS IN
CC THE MITOCHONDRION (POSSIBLY OF CATIONS). NOT ESSENTIAL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,
CC TO YEAST MRS3; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER
CC PROTEIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X56444; CAA39828.1; -
EMBL; Z28277; CAA82130.1; -
PIR; S13533; S13533.
SGD; S0001760; MRS4.
InterPro; IPR001993; Mitoch_carrier.
Pfam; PF00153; mito_carr; 3.
PROSITE; PS00215; MITOCH_CARRIER; 2.
Mitochondrion; Inner membrane; Repeat; Transport;
mRNA processing.
REPEAT 9 110 BLOCK I (APPROXIMATE).
REPEAT 111 202 BLOCK II (APPROXIMATE).
REPEAT 203 304 BLOCK III (APPROXIMATE).
TRANSMEM 23 41 POTENTIAL.

```



```
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative mitochondrial carrier C16C10.1.
GN C16C10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Lloyd C.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; Z46787; CAA86739.1; -
DR WormPep; C16C10.1; CE01489.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein; Mitochondrion; Repeat; Transmembrane;
KW Transport.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
SQ SEQUENCE 328 AA; 36743 MW; EA1E9E329A764DF6 CRC64;

Query Match 20.1%; Score 283; DB 1; Length 328;
Best Local Similarity 30.9%; Pred. No. 4.9e-19;
Matches 82; Conservative 39; Mismatches 112; Indels 32; Gaps 7;

QY 12 YRVNLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSDVI----- 64
DB 74 FRGTADAIVKTAHEGHSLSGLSPTVMALPATVFTTYDNLVSVLKKMKCCRRAPS 133

QY 65 ----HPGNSHIANAGACVATLLHDAAMNPAEVVQKQMNSPYHRVTCVRAVQNE 120
DB 134 PEKWTTPDWS--AAVAGIVARTIAVTVVSPLEMTETKMSKRLTYHEIGHVLRSMATK 191

QY 121 GAGAFYRSYTTQLTMMNVPQAIHEFTYEFLOEHFNPQRYPNPSH-----VLSGACAGAV 175
DB 192 GISSPYLGWTPMLRDIPEFGIYWAGYDLFKT--NLQRQGDHNPVSVFVSGAAAGVV 249

QY 176 AAATATPLDVCKTLNLTQESLALNSHITGHITGMASAFRTV-----YQVGGVTAYFRGVQA 231
DB 250 ASIFTHPFDVTKT--NCO-----IRIGSIDDMNKSITTVIKDMYHSGISAFSSGLVP 301

QY 232 RVVIOIPSTAIWSVYEFKYLITK 256
DB 302 RLVRVSPSCAIMSEYEFKFLFQK 326

RESULT 4
Y39_YEAST
ID Y39_YEAST STANDARD; PRT; 368 AA.
AC O03829;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative mitochondrial carrier YMR166C.
GN YMR166C OR YMR520.15C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Hunt S.; Bowman S.; Barrell B.G.; Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (Potential).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; Z49705; CAA89802.1; -
DR SGD; S0004776; YMR166C.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
SQ SEQUENCE 368 AA; 40992 MW; B583100018DF045D CRC64;

Query Match 19.6%; Score 275.5; DB 1; Length 368;
Best Local Similarity 28.6%; Pred. No. 2.8e-18;
Matches 77; Conservative 59; Mismatches 100; Indels 33; Gaps 9;

QY 11 YRVNLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSDVIHPGNN 69
DB 89 KYRNMSIAVRTIWEEGVRRGLYGGYMAALGSPFSAIFFGTYYTKRTMIEDWQI--N 146

QY 70 SHIANGACCVATLLHDAAMNPAEVVQKQM-----YHRTVCVRAVQNN 119
DB 147 DTITLSLGLDFDTSFVYVPEVLKTRLQGRNPNFPGSYNLSRLNAITVKE 206

QY 120 EGAGAFYRSYTTQLTMMNVPQAIHEFTYEFLOEHFNPQR-----YNPSSHVLSGACA 172
DB 207 EGRSLFPGYKATLARDLPFSALQFAFYEFKFLQALFKLEQKDRGELSIPNEILTGACA 266

QY 173 GAVAAAAATPLDVCKTLNLTQESLALNS-----HIT-GHITGMAS-----AFRTVYOV 219
DB 267 GGLAGIITPMDVTRVQTQPPSQSNKSYSVTHPHVTNGRPAALSNISLSLRTVYQS 326

QY 220 GGVTAIFRGVQARVYIQIPSTAIANSVYE 248
DB 327 EGVLFEGFSGVGRFVMTSVQSSIMLLLYQ 355

RESULT 5
CG69_HUMAN
ID CG69_HUMAN STANDARD; PRT; 359 AA.
AC Q9BZJ4; Q9UF66; Q9Y379; Q9P182;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial carrier protein CGI-69 (PRO2163).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SUBCELLULAR LOCATION, AND
RP TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX PubMed=11139402;
```

RA Yu X.X., Lewin D.A., Zhong A., Brush J., Schow P.W., Sherwood S.W.,
 RA Pan G., Adams S.H.;
 RT "Overexpression of the human 2-oxoglutarate carrier lowers
 RT mitochondrial membrane potential in HEK-293 cells: contrast with the
 RT unique cold-induced mitochondrial carrier CGI-69.";
 RL Biochem. J. 353:369-375(2001).
 RN [2]
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=2072150; PubMed=10810093;
 RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
 RT "Identification of novel human genes evolutionarily conserved in
 RT Caenorhabditis elegans by comparative proteomics.";
 RL Genome Res. 10:703-713(2000).
 RN [3]
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Testis;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansorge W., Boecher M., Bloeker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RN [4]
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEBO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain, and Colon;
 RA Strausberg R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN
 RP SEQUENCE OF 15-359 FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
 RA He F.;
 RT "Functional prediction of the coding sequences of 79 new genes deduced
 RT by analysis of cDNA clones from human fetal liver.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/CGI-69L (shown here) and 2;
 CC are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in many tissues. Abundant in testis
 CC and kidney.
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF317711; AAG60687.1; -
 DR EMBL; AF151827; RAD34064.1; -
 DR EMBL; AL133584; CAB63728.1; -
 DR EMBL; AK026060; BAB15341.1; -
 DR EMBL; BC001398; AAH01398.1; -
 DR EMBL; BC009330; AAH09330.1; -
 DR EMBL; AF119864; AAF69618.1; -
 DR InterPro: IPR001993; MitochCarrier.
 DR InterPro: IPR002067; MitCarrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PRINTS; PR00926; MITOCARRIER.

DR PROSITE; PS00215; MITOCH_CARRIER; 1.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Alternative splicing; Polymorphism.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT VARSPLIC 64 72 LPSSLOSTG -> W (IN ISOFORM 2).
 FT VARIANT 247 247 L-> F.
 FT CONFLICT 266 266 /FTID=VAR_012756.
 FT G -> R (IN REF. 6).
 SQ SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;
 Query Match 17.9%; Score 251.5; DB 1; Length 359;
 Best Local Similarity 27.8%; Pred. No. 4.7e-16;
 Matches 75; Conservative 43; Mismatches 121; Indels 31; Gaps 8;
 QY 7 DPAAARNVLEALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACVEKLKKTLL---60
 DB 100 DP-TRETCMDAFVKIVRHEGRTLW---SGLPATLVMTPATAIYFTAYDQLKAFLCGR 155
 QY 61 ---SDVIHPGNSHIANGAAGCVATLLHDAAMNPAEVVKQRMQMYNSPHRVYDCVRVW 117
 DB 156 ALTSLEYAP-----MVAGALARLGTVTVISPLELMRTKLQAQHVSYRELGCACVRTAV 207
 QY 118 ONEGACAFYSYTTQLTMTNVPFOAHFMFYEFLOEHFN---PORYNPNSSHVLSGACAGA 174
 DB 208 ACGGWRSLWGLMGPTALRDVPFSALYWFNYELVKSWLNGLRPKDQTSVGMFSFVAGGISGT 267
 QY 175 VAAATPTPLDVCKT---LLNTOESLALNSHTGHTGTGASAPRTVYQGVGTAYFERGVQ 230
 DB 268 VAAVLTLPFDVVKTQROVALGAEVAVNPL---HVDSTWLLLRRAESGTGKLGAGFL 324
 QY 231 ARVIQIPSTATAWSYVEFFKYLITKROEE 260
 DB 325 PRIKAAPSCAIMISTYEFKGFQRLNQD 354
 RESULT 6
 ID PET8_YEAST STANDARD; PRT; 284 AA.
 AC P38921;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Putative mitochondrial carrier protein PET8.
 GN PET8 OR YNL003C OR N2012.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GRF88;
 RX MEDLINE=94005822; PubMed=8402262;
 RA Lalo D., Stettler S., Mariotte S., Slonimski P.P., Thuriaux P.;
 RT "Two yeast chromosomes are related by a fossil duplication of their
 RT centromeric regions.";
 RL C. R. Acad. Sci., III, Sci. Vie 316:367-373(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GRF88;
 RX MEDLINE=95028151; PubMed=7941739;
 RA Lalo D., Stettler S., Mariotte S., Gendreau E., Thuriaux P.;
 RT "Organization of the centromeric region of chromosome XIV in
 RT Saccharomyces cerevisiae.";
 RL Yeast 10:523-533(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;


```
Matches 68; Conservative 41; Mismatches 110; Indels 36; Gaps 8;
QY 12 YRNVLALWRIITRGLRPMRGLNVTAGAGPAHALYFACYEKLKLTSDVI-----HP 66
Db 371 YANSDFCKKVKJRYEGFGLYKGLLPQLLGVAPAKA-----IKLTVNDVFRDKFMHK 422
QY 67 GGN-----SHIANGAAGCVATLLHDAAMNPAEAVKQRMQNSPYHRVDCVR-----AYW 117
Db 423 DGSVPLAAEILAGGAGGSGVIF---TNPLEIVKRLQVAG---EITTPRVSAHSV 474
QY 118 QNEGAGAFVRSYTTQLTNNVPQAIHFWYEFLOEHF-NPQRRYNPSSHVLGACAGAVA 176
Db 475 RDLGFFGIYKGAACFLRDIPEFSAIYFPCYAHVKASFANEDGVSPGSLLAGAIGMFA 534
QY 177 AAATPLDVKCTLLNTQESLALNSHITGTMASAFRTVYGVGTAVFRCVQARVIYQ 236
Db 535 ASLVTADVIVKRLQVAVRAGQTTI-----SGVIDCFKRLKEEGKALWKAGARVFRS 589
QY 237 IPSTAIAMSVYEFFK 251
Db 590 SPOFGVTLTYELLQ 604
RESULT 9
DNC_HUMAN
ID DNC_HUMAN STANDARD; PRT; 320 AA.
AC Q9HC21;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Mitochondrial deoxynucleotide carrier (Mitochondrial uncoupling protein 1).
OS SLC25A19 OR DNC OR MUP1.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RC TISSUE=Liver;
RA PubMed=11226231;
RT "The human mitochondrial deoxynucleotide carrier and its role in the toxicity of nucleoside antivirals."
RL Proc. Natl. Acad. Sci. U.S.A. 98:2284-2288(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Renard S., Mondesert G., Besnard F.;
RT "MUP 1, a mitochondrial uncoupling protein."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
CC -!- FUNCTION: Responsible for the uptake of deoxynucleotides into the matrix of the mitochondria. Transports all four deoxy NDPs, and, less efficiently, the corresponding dNTPs. Does not transport dNTPs, NMPs, deoxynucleosides, nucleosides, purines, or pyrimidines. Supply deoxynucleotides to the mitochondrial matrix for conversion to triphosphates and incorporation into mitochondrial DNA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.
CC -!- TISSUE SPECIFICITY: Expressed in all tissues examined except for placenta. Highest levels in colon, kidney, lung, testis, spleen, and brain.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- DISEASE: Likely to be medically important by providing the means of uptake into mitochondria of nucleoside analogs, leading to the mitochondrial impairment that underlies the toxic side effects of such drugs in the treatment of viral illnesses, including AIDS, and in cancer therapy.
CC
```

```
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ251857; CAC27560.1; -
CC EMBL; AF301616; CAC37793.1; -
CC EMBL; AF182404; AAG16903.1; -
CC EMBL; BC001075; AAH01075.1; -
CC EMBL; BC005120; AAH05120.1; -
CC MIM; 606521; -
CC InterPro; IPR001993; Mitoch_carrier.
CC InterPro; IPR002067; Mit_carrier.
CC Pfam; PF00153; mito_carr; 3.
CC PRINTS; PR00926; MITOCARRIER.
CC PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
SQ SEQUENCE 320 AA; 35511 MW; 57CE0F01D538B1BE CRC64;
Query Match 15.1%; Score 212.5; DB 1; Length 320;
Best Local Similarity 23.4%; Pred. No. 1.8e-12;
Matches 62; Conservative 54; Mismatches 116; Indels 33; Gaps 9;
QY 7 DPAARYNVLEALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKLTSDV 63
Db 54 DPSAKYHGILQASRQLQEGPTAFWKGHVPAQLISIGYGAQVQLSF-----EMLT 106
QY 64 IHGGG-----NSH- IANGAAGCVATLLHDAAMNPAEAVKQRMQNSPY--YHRVTD 113
Db 107 VHRGVSVDAREFSVHFVCGGLACMATL---TVHPVDVLRTRFAAQGPKVYNTLRHAV 162
QY 114 RAWNOGAGAFVRSYTTQLTNNVPQAIHFWYEFLOEHF-----PQRRYNPSSHVL 168
Db 163 GTMYRSEGPQVFKGLAPTLIAFPYAGLQFCYSLSKHLKXWAPAECKKNENLQNLIC 222
QY 169 GACAGAVAAAATPLDVKCTLLNT---QESLALNSHITGTMASAFRTVYGVGTAY 225
Db 223 GSGAGVISKTLTYPLDLFKRLQVGVGFHARAAGVRRY-RGLMDCAKQVLQKEGALGF 281
QY 226 FRGVQARVIYQIPSTAIAMSVYEFF 250
Db 282 FKGLSPLSLKAAALSTGFMEFFSYEFF 306
RESULT 10
CMCL_DROME
ID CMCL_DROME STANDARD; PRT; 695 AA.
AC Q9VA73; Q9VA72; Q9VA74; Q9U5V8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-binding mitochondrial carrier Atalar1.
GN ATALAR1 OR CG2139.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Ovary;
RX MEDLINE=20115110; PubMed=10642534;
```

RA Del Arco A., Auado M., Satrutegui J.;
 RT "Characterization of a second member of the subfamily of calcium-
 binding mitochondrial carriers expressed in human non-excitable
 tissues.";
 RL Biochem. J. 345:725-732(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Sridaling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Sheng H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN; MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; MAY BE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y18197; CAB62169.1; -
 CC EMBL: AE003774; AAF57048.1; -
 CC EMBL: AE003774; AAF57049.1; -
 CC EMBL: AE003774; AAF57050.2; -
 CC FlyBase: FBgn0028646; aralar1.
 CC InterPro: IPR002048; EF-hand.
 CC InterPro: IPR002067; Mit_carrier.
 CC InterPro: IPR001993; Mitoch_carrier.

DR Pfam: PF00036; ehand; 3.
 DR Pfam: PF00153; mito_carr; 3.
 DR PRINTS: PR00926; MITOCARRIER.
 DR SMART: SM00054; Eph; 3.
 DR PROSITE: PS00018; EF_HAND; 1.
 DR PROSITE: PS0015; MITOCH_CARRIER; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Calcium-binding; Alternative splicing.
 FT TRANSMEM 346 363 1 (POTENTIAL).
 FT TRANSMEM 406 425 2 (POTENTIAL).
 FT TRANSMEM 449 462 3 (POTENTIAL).
 FT TRANSMEM 498 517 4 (POTENTIAL).
 FT TRANSMEM 537 554 5 (POTENTIAL).
 FT TRANSMEM 594 613 6 (POTENTIAL).
 FT CA_BIND 84 95 EF-HAND 1.
 FT CA_BIND 118 129 EF-HAND 2.
 FT DOMAIN 157 168 ANCESTRAL CALCIUM SITE 3.
 FT CA_BIND 189 200 EF-HAND 4.
 FT DOMAIN 225 227 3 APPROXIMATE TANDEM REPEATS.
 FT REPEAT 225 227 1.
 FT REPEAT 435 524 2.
 FT REPEAT 525 627 3.
 FT VARSPPLIC 1 52 MISSING (IN ISOFORM 3).
 FT VARSPPLIC 3 22 MHIPFPNWIPLPVARCOE -> LTKSLPN (IN
 FT ISOFORM 2).
 FT CONFLICT 356 357 GA -> AP (IN REF. 1).
 FT CONFLICT 556 556 V -> A (IN REF. 1).
 FT CONFLICT 653 653 V -> L (IN REF. 1).
 FT SEQUENCE 695 AA; 76753 MW; 8EC93D92031F5B77 CRC64;
 SQ
 Query Match 14.9%; Score 210; DB 1; Length 695;
 Best Local Similarity 28.0%; Pred. No. 7.8e-12;
 Matches 72; Conservative 36; Mismatches 113; Indels 36; Gaps 8;
 QY 12 YRNVLEALWRIIRTEGLWPMRGLNVTATGAGPAHALYFACYEKLKTLSDVTH----- 65
 DB 384 YRNSWDFCKVVRHGFGLGLPQLMGVAPEKA-----IKLVNDLVDRDKLTK 435
 QY 66 ----PGNSHIANGAAGCVATLLHDAAMPAEVVVKRMQ----MYSNPYHRVTDVAVV 117
 DB 436 KGNIPWAEVLAGGAGASQVVF----TNPLEIVKIRLOVAGEIASGSKIRANSVVREL- 490
 QY 118 QNEGAGAFYSYTTQTMNVFQAIHFMYEFLQEHNFQRRYN-PSSHVLSGACAGAVA 176
 DB 491 ---GLFLYKRGARCLRDVPFSAIYFPTYAHTKAMADKDGYNHPLTLAAGAIAGVPA 547
 QY 177 AATTPDLVCKTLNLTQESIALNSHLTGHTGMASAFRTVYQVGVTA YFRGVOARVIYQ 236
 DB 548 ASLVTADVIKTRL---QVVARSGQTT--YTGVDATKIKMAEEGPRAFWKGTAAARVFRS 602
 QY 237 IPSTAIWSVYEFKYL 253
 DB 603 SPQFGVTLTYELLQRL 619
 RESULT 11
 BTL_MAIZE
 ID BTL_MAIZE STANDARD; PRT; 436 AA.
 AC P29518;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Brittle-1 protein, chloroplast precursor.
 GN BTL.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93005685; PubMed=1668652;

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94203187; PubMed=7908717;
 RA Fernandez M., Fernandez E., Rodicio R.;
 RT "ACR1, a gene encoding a protein related to mitochondrial carriers,
 is essential for acetyl-CoA synthetase activity in Saccharomycetes
 cerevisiae."; Mol. Gen. Genet. 242:727-735(1994).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP Ramezani Rad M., Kirchthath L., Hollenberg C.P.;
 RA Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN CHARACTERIZATION.
 RX MEDLINE=98055465; PubMed=9395087;
 RA Palmieri L., Lasorsa F.M., De Palma A., Palmieri F., Runswick M.J.,
 RA Walker J.E.;
 RT "Identification of the yeast ACR1 gene product as a succinate-fumarate
 transporter essential for growth on ethanol or acetate."; FEBS Lett. 417:114-118(1997).
 RL [4]
 CC -!- FUNCTION: Transports cytoplasmic succinate, derived from
 isocitrate by the action of isocitrate lyase in the cytosol, into
 the mitochondrial matrix in exchange for fumarate.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial.
 CC -!- INDUCTION: BY ETHANOL OR ACETATE AS SOLE CARBON SOURCES. REPRESSED
 BY GLUCOSE.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z25485; CA80973.1; -;
 DR EMBL; Z49595; CA89624.1; -;
 DR PIR; S36407; S36407.
 DR PIR; S43280; S43280.
 DR SGD; S0003856; SFC1.
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE; PS00215; MITOCH_CARRIER; 3.
 KW Transmembrane; transport; Repeat.
 KW CONFLICT 300 322 VREHLEMLGIFKNDTPKPLK -> RKGAFOKIWVYSRR
 FT MTHQSOSH (IN REF. 1).
 SQ SEQUENCE 322 AA; 35340 MW; EC29718A0F5011A7 CRC64;

 Query Match 14.78; Score 206; DB 1; Length 322;
 Best Local Similarity 26.68; Pred. No. 7.2e-12;
 Matches 69; Conservative 35; Mismatches 107; Indels 48; Gaps 10;

 Qy 22 IIRTEGLWRPRLNVTAGAGPAHALYFACYEKLKLTLDVHPGNSHANGAA 77
 Db 62 IYQEGFALYKGLGAVVIGIPIPKMAIRFSEFYRTLILVNKESGIVST-GNTEFVAGVA 120
 Qy 78 GCVATLHDAAMNPAEVVKQRMQ-----MYNSPYHRVDCVRAVWQNEGAGA 124
 Db 121 GITEAVL---VYNMEVVKIRLQAOHLTPSPNAGPKYNNIAHAAYITVK-----EEGVSA 173
 Qy 125 FYRSYTTOLTNNVPFOATHFTY----EFLQEHFNPRRYNPSSHV-LSGACAGAVAAAA 179
 Db 174 LYRGVSLTAARQATNOGANFTVYSKLKEFLQNYH--QMDVLPSTWETSIGLISGAIGPFS 231
 Qy 180 TTPDVKCTLNTQESALNSHITGHTGMASAFRTVYQVG-----GVTAIFRGVQAR 232
 Db *232 NAPLDTITKLRQKDSISLEQO-----SGMKIITIGAOQLKBEGRFALYKGTIPR 282

Qy 233 VIYQIPSTAIANSVYEFEK 251
 Db 283 VNRVAPGQAVTFTVYEVYR 301

 RESULT 14
 YAD8_SCHPO STANDARD; PRT; 271 AA.
 AC Q09834;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE putative mitochondrial carrier C4G8.08.
 GN SPAC48.08.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 inner membrane (Potential).
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z56276; CA91209.1; -;
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE; PS00215; MITOCH_CARRIER; 1.
 KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
 KW Transmembrane; transport.
 FT TRANSMEM 49 69 POTENTIAL.
 FT TRANSMEM 84 104 POTENTIAL.
 FT TRANSMEM 138 158 POTENTIAL.
 FT TRANSMEM 170 190 POTENTIAL.
 SQ SEQUENCE 271 AA; 30002 MW; EF330C8CC8BDB6C4 CRC64;

 Query Match 14.58; Score 204; DB 1; Length 271;
 Best Local Similarity 27.08; Pred. No. 8.9e-12;
 Matches 61; Conservative 41; Mismatches 110; Indels 14; Gaps 5;

 Qy 33 RGLNVATGAGPAHALYFACYEKLKLTLDVHPGNSHANGAAAGCVATLLHDAAMNPA 92
 Db 48 RGLPISLTLTPATCLYLSTVTEAKRRFKPSV---GEGAILYSGMTAEVVSFVWTPPL 104
 Qy 93 EVVKQRMQMYNSPYHRVDCVRAVWQNEGAFYRSYTTQLTNNVPFOAIHFMTYEFLOE 152
 Db 105 EVIKARTQI--SQKGSVINTISTLARSEGLKGFYRGYWMGVAIYLPFTVSWVCYESKK 162
 Qy 153 HFNQRRYNPSSHVLSGACAG---AVAAAATPDLVKCTLNTQESALNSHITG----H 205
 Db 163 YLQKSNWDIS--VIAPICSLGTVVATTISTPLDVKTRVQVATSSAMRKAETGLQAEK 220
 Qy 206 ITGMASAFRTVYQVGVTAYFRGVQARVIYQIIPSTAIANSVYEFEK 251
 Db 221 ELGILEIAKLFLSKHVGKGTGRGLFTRMCIYMPSGMISMSVFESEK 266

 RESULT 15


```
ODC2_YEAST          STANDARD;          PRT;          307 AA.
ID   ODC2_YEAST
AC   Q99297;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Mitochondrial 2-oxodicarboxylate carrier 2.
GN   ODC2 OR YOR222W OR YOR50-12.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RX   MEDLINE=9643797; PubMed=8840505;
RA   Galisson F., Dujon B.;
RT   "Sequence and analysis of a 33 kb fragment from the right arm of
RL   chromosome XV of the yeast Saccharomyces cerevisiae.";
RL   Yeast 12:877-885(1996).
RN   [2]
RP   CHARACTERIZATION.
RX   PubMed=11013234;
RA   Palmieri L., Adimi G., Runswick M.J., Fearnley I.M., Palmieri F.,
RA   Walker J.E.;
RT   "Identification in Saccharomyces cerevisiae of two isoforms of a novel
RL   J. Biol. Chem. 276:1916-1922(2001).
CC   -1- FUNCTION: Transports C5-C7 oxodicarboxylates across the inner
CC   membranes of mitochondria. Can transport 2-oxoadipate, 2-
CC   oxoglutarate, adipate, glutarate, 2-oxopimelate, oxaloacetate,
CC   citrate and malate. The main physiological role is probably to
CC   supply 2-oxoadipate and 2-oxoglutarate from the mitochondrial
CC   matrix to the cytosol where they are used in the biosynthesis of
CC   lysine and glutamate, respectively, and in lysine catabolism.
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC   inner membrane.
CC   -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC   -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@sib-sib.ch).
CC   -----
DR   EMBL; Z75130; CAA99440.1; -.
DR   EMBL; X92441; CAA63185.1; -.
DR   SGD; S0005748; ODC2.
DR   InterPro; IPR002067; Mit_carrier.
DR   InterPro; IPR001993; Mitoch_carrier.
DR   Pfam; PF00153; mito_carr; 3.
DR   PRINTS; PR00926; MITOCARRIER.
DR   PROSITE; PS00215; MITOCH_CARRIER; 3.
KW   Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT   TRANSMEM 10..30  POTENTIAL.
FT   TRANSMEM 122..142  POTENTIAL.
FT   TRANSMEM 215..235  POTENTIAL.
FT   TRANSMEM 280..300  POTENTIAL.
SQ   SEQUENCE 307 AA; 34007 MW; 4089082A64DBA97C CRC64;

Query Match          14.2%; Score 199.5; DB 1; Length 307;
Best Local Similarity 24.1%; Pred. No. 2.7e-11;
Matches 62; Conservative 51; Mismatches 119; Indels 25; Gaps 6;

QY  11 RYRNVLALWRIITRGLWRPMGLNVTAGPAHALYFACYEKLKTKLSDVIHPGNS 70
DB  58 RYNGVIDCLKKIVKKEGFSKLYGISPMLMEAPKATKFCNCDQYOKIFKNLFTNETT 117
QY  71 HIANGAAGCVATLLHDAAMNPAEVVVKORMQYNPSYHRVTDVCRVAVWQNEGAGAFYSYT 130
```

```
Db 118 QKISIAAGASAGMTEAAVIVPPELIKIRMQDVKSSYLGPMDCCLKTKIKNEGIMGLYKIE 177
QY 131 TOLTWNYPFOAIHF-MTYEFLQEHFNPPRRYNP-----SSHVLSCACAGAAAAAT 180
Db 178 STMWRNALWNGGYFGVIYQV-----RNSMPVAKTKGQKTRNDLIAGAI GGIVGTMLN 229
QY 181 TPLDVCKTLLNTQESLALNSHITGHTGMSAFRTVYQVGWGTAYFRGVQARVIYQIPST 240
Db 230 TPDFDVKSRIQSVD--AVSSAVKKNWCLPSLL-VIYREEGFRALYKGFVPKVCLELAPGG 286
QY 241 AIAWSVY-----EFFKYL 253
Db 287 SLMLVWFETGMNFFRDL 303
```

Search completed: August 27, 2002, 04:16:48
Job time: 1240 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1406	100.0	364	4	Q96A46	Q96A46 homo sapien
2	1007.5	71.7	338	11	Q92OQ8	Q92OQ8 mus musculu
3	937	66.6	177	4	Q969G8	Q969G8 homo sapien
4	910	64.7	176	4	Q9H2J3	Q9H2J3 homo sapien
5	848.5	60.3	347	4	Q9NY22	Q9NY22 homo sapien
6	725.5	51.6	187	4	Q9P0J2	Q9P0J2 homo sapien
7	722.5	51.4	187	11	Q9CGG7	Q9CGG7 mus musculu
8	677	48.2	312	5	Q23125	Q23125 caenorhabdi
9	657.5	46.8	379	5	Q9VAV3	Q9VAV3 drosophila
10	643	45.7	380	5	Q94638	Q94638 onchocerca
11	594	42.2	380	5	Q9NHV6	Q9NHV6 drosophila
12	573	40.8	301	5	Q94634	Q94634 onchocerca
13	467	33.2	303	3	Q14281	Q14281 schizosacch
14	431.5	30.7	781	10	Q9LMW6	Q9LMW6 arabidopsi
15	429.5	30.5	289	10	Q82049	Q82049 ribes nigru
16	414	29.4	331	10	Q64731	Q64731 arabidopsi


```

ID Q9NYZ2 PRELIMINARY; PRT; 347 AA.
AC Q9NYZ2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE HT015 PROTEIN.
GN HT015.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=ADRENAL GLAND;
RC Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RA "A novel gene expressed in human hypothalamus.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223466; AAF64141.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN.2.
SQ SEQUENCE 347 AA; 37828 MW; 700DE61B230E001E CRC64;

Query Match 60.3%; Score 848.5; DB 4; Length 347;
Best Local Similarity 68.4%; Pred. No. 7.5e-75;
Matches 154; Conservative 34; Mismatches 36; Indels 1; Gaps 1;

QY 1 MQSLQPPAARYRNVLAEALNRIRTEGLWRPRLGNTATGAPAHALYFACYEKLKRTL 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 MQSLSPDKAQYTSIYGALKIMTEGFWRLGYNVIMGAPAHAMFYACYENMKRTL 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 SDVTHPCGNSHIANGACGVATLLHDAAMPAEVVKORMQYNPSYHRVTDVRAVQNE 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 NDVFHHGNSHLANGAGSATLLHDAVMPAEVVKORLQYNSQHSATSIETVWRT 192

QY 121 GAGAFYSYTTQTMNYPFOAIHPTMEYFLOEHFNPNRRYNSPSSHVLSGACAGAVAAAT 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 GLGAFYSYTTQTMNYPFOAIHPTMEYFLOEHFNPNRRYNSPSSHVLSGACAGAVAAAT 252

QY 181 TPLDVCKTLTQBSLALN-SHITGHTITGMAFRVYQGVGTA 224
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 TPLDVCKTLTQBSLALN-SHITGHTITGMAFRVYQGVGTA 297

RESULT 6
Q9POJ2 PRELIMINARY; PRT; 187 AA.
AC Q9POJ2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=ADRENAL GLAND;
RC Huang Q., Yu Y., Huang C., Jiang C., Ren S., Zhou J., Gu W., Xu S.,
RA Wang Y., Fu G., Chen Z., Han Z.;
RA "A novel gene expressed in human adrenal gland.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF155660; AAF67479.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 2.
DR PRINTS; PR00926; MITOCH_CARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN.1.
SQ SEQUENCE 187 AA; 21125 MW; 26F760526F7DE21B CRC64;

```

```

Query Match 51.6%; Score 725.5; DB 4; Length 187;
Best Local Similarity 71.9%; Pred. No. 3.7e-63;
Matches 133; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY 80 VATLLHDAAMNPAEVVKORMQYNPSYHRVTDVRAVQNEGAGAFYSYTTQTMNYPF 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MATLLHDAVMPAEVVKORLQYNSQHSATSIETVWRTGEGAFYSYTTQTMNYPF 60

QY 140 QAIHPTMEYFLOEHFNPNRRYNSPSSHVLSGACAGAVAAATTPLDVCKTLTQBSLALN 199
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 QSIHPTMEYFLOEHFNPNRRYNSPSSHVLSGACAGAVAAATTPLDVCKTLTQBSLALN 120

QY 200 -SHITGHTITGMAFRVYQGVGTAFRVYQGVGTAFRVYQGVGTAFRVYQGVGTA 258
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 LANISGRSLGMAFRVYQGLAGYFKGIQARVITQMPSTALSWSEYFVKFLIKRQ 180

QY 259 EEWRA 263
   |||
Db 181 LENRA 185

RESULT 7
Q9CQG7 PRELIMINARY; PRT; 187 AA.
AC Q9CQG7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE 1700020E22RIK PROTEIN.
GN 1700020E22RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK015770; BAB29969.1; -.
DR EMBL; AK006155; BAB24436.1; -.
DR MGD; MGI:1914962; 1700020E22Rik.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 2.
DR PRINTS; PR00926; MITOCH_CARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN.1.
SQ SEQUENCE 187 AA; 21194 MW; 8FB23583A8DD5E90 CRC64;

Query Match 51.4%; Score 722.5; DB 11; Length 187;
Best Local Similarity 71.2%; Pred. No. 7.4e-63;
Matches 131; Conservative 30; Mismatches 22; Indels 1; Gaps 1;

```

```
QY 80 VATLLHDAAMNPAEVYKQRMQNSYHRYHVTCVRAVWQNEGAGAFYRSTYTTQLTMTNVPF 139
Db 1 MATLLHDAVMPNPAEVYKQRLQWNSQSAFSCIRTVWRTGELGAFYRSTYTTQLTMTNIPF 60
QY 140 OATHMTYEFLOHFNPPORRYNPSSHVLSCACAGAAVAAATPLDVCCKTLNLTQESLALN 199
Db 61 QSHFTYEFLOHFNPPORRYNPSSHVLSCACAGAAVAAATPLDVCCKTLNLTQENMALN 120
QY 200 -SHITGHITGMAFAFTYVQGVGTAFYRGVQARVYIPIPTAIWSVVEFFKYLITKQ 258
Db 121 LANVSGRLSGMANAFYVQLNGLAGYFKGIQARVYIOMPSTAIWSVVEFFKYLITKQ 180
QY 259 EEMR 262
Db 181 LENR 184
RESULT 8
Q23125 PRELIMINARY; PRT; 312 AA.
AC Q23125;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE W02B12.9 PROTEIN.
GN W02B12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swinburne J., Ainscough R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RT none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z66521; CAA91399.1; -.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;
Query Match 48.2%; Score 677; DB 5; Length 312;
Best Local Similarity 50.6%; Pred. NO. 4.2e-58;
Matches 133; Conservative 47; Mismatches 73; Indels 10; Gaps 3;
QY 1 MOSLOPDPAAARYNNVLEALWRIITRTGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
Db 45 MOSLCPCTKPTPTVHSLMSLVKREGWLPURGVNAVAGSMPAHALYFTYERKMGYL 104
QY 61 SDVIHPGNSH-IANGAGCVATLLHDAAMNPAEVYKQRMQNSYHRYHVTCVRAVWQ 119
Db 105 TG--NSAGHSNTLAYGAGVATLLHDAIMNPAEVYKQRMQMAFSPYGSLSLECARCVNR 162
QY 120 ECAGAFYRSYTTQLTMTNVPFQAIHMTVEFLOEHFNPPORRYNPSSHVLSCACAGAAVAAA 179
Db 163 EGVAFAFYRSYTTQLTMTNVPFQAIHMTVEFLOEHFNPPORRYNPSSHVLSCAGLAGLAAL 222
QY 180 TPLDVCCKTLNLTQESLALN-----HITGHITGMAFAFTYVQGVGTAFYRGVQAR 232
Db 223 TTPMDCVKTVLTQAAADPANRIFLQARYRYRGISDAVTVIYSQGLSFGSGLQAR 282
QY 233 VTYIPTSTAIWSVVEFFKYLIT 255
Db 283 VIFQVPATALWSVVEFFKFLMS 305
```

```
RESULT 9
Q9VAY3 PRELIMINARY; PRT; 379 AA.
AC Q9VAY3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CG4963 PROTEIN (GH09840P).
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadelis E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Gual J., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL: AF003763; AAF56764.1; -.
DR EMBL: AY060268; AAL25307.1; -.
DR FlyBase: FBgn0039561; CG4963.
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR002067; Mit_carrier.
```

```
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 46.8%; Score 657.5; DB 5; Length 379;
Best Local Similarity 51.2%; Pred. No. 4.5e-56;
Matches 130; Conservative 35; Mismatches 76; Indels 13; Gaps 3;

QY 1 MOSLOPPAAPAARNVLEALRIITTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
Db 42 MQSL--SPPTKNNIVSTLTREGLLRPIRGASAVILGAGPAHSLYFAAYEMTKELT 99
QY 61 SDVIHPGGNSHIANGAGCVATLLHDAAMNPAEVVKORMQYNSPYHRVDCVRVWQNE 120
Db 100 AKTTSVRLNLYVISA---VATLIHDAISSPTDVIKORMQYNSPYTSVSCVRDIYKR 156
QY 121 GAGAFYRSYTTQLTMMNVPFOAIHFMTYEFLOEHFNPPORRYNPSSHVLGACAGAVAAAAT 180
Db 157 GFRAFYRAYGTQLVMNLPYQTIHFTYEFQKNKLERKYNPPVHMAAGAAAGACAAAVT 216
QY 181 TPLDVCKTLTNTQESLALNSHITGTMASAFRTYVQGVGTAYFRGVQARVIYQIPST 240
Db 217 TPLDVTKTLTNTQE-----TGLTRGMIEASRKIYHMAAGPLGFFRGTTARVLYSMPAT 268
QY 241 AIAWSVYEFFKYLI 254
Db 269 AICWSYIEFFKYFL 282

RESULT 10
Q94638 PRELIMINARY; PRT; 303 AA.
AC Q94638;
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45998; AAB19037.1; -.
DR InterPro; IPR001993; Mitoch_Carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 303 AA; 33861 MW; 224BFA547E5D617C CRC64;

Query Match 45.7%; Score 643; DB 5; Length 303;
Best Local Similarity 46.4%; Pred. No. 8.7e-55;
Matches 122; Conservative 48; Mismatches 83; Indels 10; Gaps 2;

QY 1 MOSLOPPAAPAARNVLEALRIITTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
Db 42 LOSLCPCPETSCPTAMHSLMSWKREGLLRSLKGVNANVLGTIPARAFYTVTVENSAYL 101
QY 61 SDVIHPGGNSHIANGAGCVATLLHDAAMNPAEVVKORMQYNSPYHRVDCVRVWQNE 120
Db 102 LN--NPRVNSVSVAISGALATYIHDAVMNPAEVVKORMQYNSPYGNSLECIRCIYRE 159
QY 121 GAGAFYRSYTTQLTMMNVPFOAIHFMTYEFLOEHFNPPORRYNPSSHVLGACAGAVAAAAT 180
Db 160 GLRAFYRSYTTQLTMMNVPYQCTHEMIYEQNLNPHHDYNPSSHVLVSGGIAGGIAAAT 219

Query Match 42.2%; Score 594; DB 5; Length 380;
Best Local Similarity 47.5%; Pred. No. 7.5e-50;
Matches 121; Conservative 35; Mismatches 85; Indels 14; Gaps 4;

QY 1 MOSLOPPAAPAARNVLEALRIITTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
Db 42 MQSL--SPPTKNNIVSTLTREGLLRPIRGASAVILGAGPAHSLYFAAYEMTKELT 99
QY 61 SDVIHPGGNSHIANGAGCVATLLHDAAMNPAEVVKOR-MQMYNSPYHRVDCVRVWQNE 119
Db 100 AKTTSVRLNLYVISA---VATLIHDAISSPTDVIKAYADCTTRPHTPVVSCVRDIYKR 156
QY 120 EGAGAFYRSYTTQLTMMNVPFOAIHFMTYEFLOEHFNPPORRYNPSSHVLGACAGAVAAA 179
Db 157 EGKAFYRAYGTQLVMNLPYQTIHFTYEFQKNKLERKYNPPVHMAAGAAAGACAAAV 216
QY 180 TPLDVCKTLTNTQESLALNSHITGTMASAFRTYVQGVGTAYFRGVQARVIYQIP 239
Db 217 TPLDVTKTLTNTQE-----TGLTRGMIEASRKIYHMAAGPLGFFRGTTARVLYSMPA 268
QY 240 TAIWSVYEFFKYLI 254
Db 269 TALCWSYIEFFKYFL 283

RESULT 12
Q94634 PRELIMINARY; PRT; 301 AA.
AC Q94634;
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Schmid K.J., Tautz D.;
RT "A screen for rapidly evolving genes from Drosophila.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217402; AAF73387.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_Carrier.
DR InterPro; IPR002067; Mit_Carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 41844 MW; 4C9AA524B97F8C6C CRC64;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:55:59 ; Search time 138.54 Seconds
(without alignments)
75.364 Million cell updates/sec

Title: US-09-870-113-10

Perfect score: 502

Sequence: 1 MQSLQPDPAARYNRVLEALW.....GAGCVATLLHDAAMNPAEG 94

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496	98.8	289	21	Human uncoupling p
2	496	98.8	289	22	Human transport pr
3	496	98.8	318	22	Human polypeptide
4	467	93.0	268	22	Human polypeptide
5	352	70.1	299	22	Human protein SEQ
6	351	69.9	272	21	Human ORFX ORF2744
7	351	69.9	331	22	Human mitochondria
8	351	69.9	366	22	Human bone marrow
9	322	64.1	677	22	Novel human secret
10	276	55.0	155	21	Human ORFX ORF2730
11	171.5	34.2	379	22	Drosophila melanog

12	165.5	33.0	59	22	AAM85891	Human immune/haema
13	130.5	26.0	289	18	AAW17054	Blackcurrant RIB7
14	130	25.9	278	21	AAG22079	Arabidopsis thalia
15	130	25.9	278	21	AAG43094	Arabidopsis thalia
16	130	25.9	290	21	AAG22078	Arabidopsis thalia
17	130	25.9	290	21	AAG43093	Arabidopsis thalia
18	130	25.9	331	21	AAG22077	Arabidopsis thalia
19	130	25.9	331	21	AAG43092	Arabidopsis thalia
20	121	24.1	130	22	AM00825	Human bone marrow
21	99.5	19.8	413	22	ABB69108	Drosophila melanog
22	95.5	19.0	310	22	AB93001	Human protein sequ
23	95.5	19.0	311	21	AB50382	Human uncoupling p
24	95.5	19.0	311	22	AM39173	Human polypeptide
25	95.5	19.0	374	22	AM40959	Human polypeptide
26	95.5	19.0	374	22	AM40960	Human polypeptide
27	94.5	18.8	237	21	AAG19645	Arabidopsis thalia
28	94.5	18.8	237	21	AAG50519	Arabidopsis thalia
29	94.5	18.8	241	21	AAG07172	Arabidopsis thalia
30	94.5	18.8	284	21	AAG19644	Arabidopsis thalia
31	94.5	18.8	284	21	AAG50518	Arabidopsis thalia
32	94.5	18.8	289	21	AAG07171	Arabidopsis thalia
33	94.5	18.8	296	21	AAG07170	Arabidopsis thalia
34	94.5	18.8	313	21	AAG19643	Arabidopsis thalia
35	94.5	18.8	313	21	AAG50517	Arabidopsis thalia
36	89.5	17.8	351	21	AB50378	Human uncoupling p
37	89.5	17.8	365	21	AB42783	Human ORFX ORF2547
38	89	17.7	320	21	AB50379	Human uncoupling p
39	89	17.7	320	22	AM79071	Human protein SEQ
40	89	17.7	320	22	AM93892	Human polypeptide
41	89	17.7	329	22	AM42411	Human polypeptide
42	89	17.7	329	22	AM42413	Human polypeptide
43	89	17.7	335	21	AB51795	Human secreted pro
44	88	17.5	358	22	ABB71142	Drosophila melanog
45	87	17.3	289	20	AA31936	Wheat brittle-1 pa

ALIGNMENTS

RESULT 1

AA50383
ID AA50383 standard; Protein; 289 AA.

AC AA50383;

DT 12-MAR-2001 (first entry)

DE Human uncoupling protein #6.

DE Human uncoupling protein; immunosuppressive; antiarthritic;

KW antirheumatic; antiproliferative; cardiant; vasotropic;

KW cerebroprotective; neuroprotective, antibacterial; ophthalmological;

KW gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;

KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;

KW infertility.

OS Homo sapiens.

PN WO2000061614-A2.

PD 19-OCT-2000.

PF 06-APR-2000; 2000WO-US09534.

PR 09-APR-1999; 99US-0128701.

PR 08-JUL-1999; 99US-0142821.

PR 18-AUG-1999; 99US-0149448.

PR 12-NOV-1999; 99US-0164751.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;

DR WPI; 2000-656322/63.
 DR N-PSDB; AAC90457.
 XX
 PT Uncoupling proteins and nucleic acid sequences encoding them, useful
 PT for detecting, preventing and treating proliferative, neurological,
 PT immune system, cardiovascular and gastrointestinal disorders -
 XX
 PS Claim 11; Page 323-324; 343pp; English.
 XX
 CC The present sequence is a human uncoupling protein. The nucleotide
 CC sequences encoding the uncoupling proteins may be used for
 CC the detection of various disorders such as cancer, for chromosome
 CC identification, as chromosome markers and for numerous other diagnostic
 CC or research purposes. The uncoupling protein encoded by the nucleotide
 CC sequences may be used to treat disorders such as neural, immune,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders, wounds, infectious diseases,
 CC thrombosis, arthritis, and infertility.
 XX
 SQ Sequence 289 AA;
 Query Match 98.8%; Score 496; DB 21; Length 289;
 Best Local Similarity 100.0%; Pred. NO. 4.4e-54;
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MQSLQDPAPARYRNVLALWRIIRTEGLWRPGRGLNVTATGAGPAHALYFACYEKLKTL 60
 Db 25 mqlqpdpaparyrnvlealwriirteglwrpgrglnvtatgagpahalyfacyeklkttl 84
 Qy 61 SDVIHPGNSHIANGAGCVATLLHDAAMNPAE 93
 Db 85 sdvihpgnshiangaagcvatllhdaampae 117
 RESULT 2
 AAB60113
 ID AAB60113 standard; Protein; 289 AA.
 XX
 AC AAB60113;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE Human transport protein TPPT-33.
 XX
 KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
 KW neurological disorder; cardiovascular disorder; reproductive disorder;
 KW immune disorder; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200078953-A2.
 XX
 PD 28-DEC-2000.
 XX
 PF 16-JUN-2000; 2000WO-US16668.
 XX
 PR 17-JUN-1999; 99US-0139923.
 PR 10-AUG-1999; 99US-0148177.
 PR 18-AUG-1999; 99US-0149357.
 PR 28-OCT-1999; 99US-0162287.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
 PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
 XX
 DR WPI; 2001-041424/05.
 DR N-PSDB; AAF27733.
 XX
 PT Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -

XX Claim 2; Page 133-134; 165pp; English.
 XX
 CC The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPPTS). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.
 XX
 SQ Sequence 289 AA;
 Query Match 98.8%; Score 496; DB 22; Length 289;
 Best Local Similarity 100.0%; Pred. NO. 4.4e-54;
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MQSLQDPAPARYRNVLALWRIIRTEGLWRPGRGLNVTATGAGPAHALYFACYEKLKTL 60
 Db 25 mqlqpdpaparyrnvlealwriirteglwrpgrglnvtatgagpahalyfacyeklkttl 84
 Qy 61 SDVIHPGNSHIANGAGCVATLLHDAAMNPAE 93
 Db 85 sdvihpgnshiangaagcvatllhdaampae 117
 RESULT 3
 AAM41505
 ID AAM41505 standard; Protein; 318 AA.
 XX
 AC AAM41505;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 6436.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI60661.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 2; SEQ ID NO 6436; 10078pp; English.
 XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 318 AA;

Query Match 98.8%; Score 496; DB 22; Length 318;
Best Local Similarity 100.0%; Pred. No. 4.9e-54;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQSLQDPDAARYNVLEALWRIIRTEGLWRPMLRGVLTATGAGPAHALYFACYEKLKKTLL 60
DB 54 mqsldpdaarynvlealwriirteglwrpmlrgvltatgagpahalyfacyekikktll 113
QY 61 SDVIHPGNGSHANGAAGCVATLLHDAAMPAAE 93
DB 114 sdvihpgngshangaagcvatllhdaampae 146

RESULT 4
AAM39719
ID AAM39719 standard; Protein; 268 AA.

XX AAM39719;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2864.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI: 2001-442253/47.
DR N-PSDB; AAI58875.

XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -

XX Example 4; SEQ ID NO 2864; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AA42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.

XX Sequence 268 AA;

Query Match 93.0%; Score 467; DB 22; Length 268;
Best Local Similarity 94.6%; Pred. No. 1.8e-50;
Matches 88; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MQSLQDPDAARYNVLEALWRIIRTEGLWRPMLRGVLTATGAGPAHALYFACYEKLKKTLL 60
DB 4 mhsldspagrynvleviirteglwrpmlrgvltatgawpahalyfacyekikktll 63
QY 61 SDVIHPGNGSHANGAAGCVATLLHDAAMPAAE 93
DB 64 sdvihpgngshangaagcvatllhdaampae 96

RESULT 5
AAM79039
ID AAM79039 standard; Protein; 299 AA.

XX AAM79039;

XX 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1701.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0663561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.


```

XX OS Homo sapiens.
XX PN CN1269409-A.
XX PD 11-OCT-2000.
XX PF 17-MAR-2000; 2000CN-0114958.
XX PR 17-MAR-2000; 2000CN-0114958.
XX PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
XX PI Zhang X, Gao X, Xiao H;
XX DR WPI: 2001-050544/07.
XX DR N-PSDB; AAF59920.
XX PT New human mitochondrion solute carrier protein and its nucleic acid -
XX PS Claim 4; Page 20; 21pp; Chinese.
XX CC The invention relates to a novel human mitochondrial solute carrier
CC protein, hMSC-o (AAF60658), and cDNA encoding it (AAF59920). hMSC-o is
CC expressed in normal human hypothalamus tissue. The invention also relates
CC to the preparation of hMSC-o proteins and nucleic acids, and the
CC detection of hMSC-o proteins and nucleic acids in a sample. The present
CC sequence represents hMSC-o.
XX SQ Sequence 331 AA;

Query Match 69.9%; Score 351; DB 22; Length 331;
Best Local Similarity 67.7%; Pred. No. 1e-35;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLQPDPAARYNNLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
Db ||||| ||| : : || : ||||| ||||| ||||| ||||| ||||| : |||
57 mqslspdkaqytsiygalkkimrtgfwprlgvnmimgagpahamyfacyenmkrtl 116
QY 61 SDVIHPGGNSHTANGAAGCVATILLHDAANPAE 93
Db : ||| ||||| ||| || : ||||| |||||
117 ndvfhhqgnshlangiagsmatilhdavnmpae 149

RESULT 8
AAM00938
ID AAM00938 standard; Protein; 366 AA.
XX AC AAM00938;
XX DT 01-OCT-2001 (first entry)
XX DE Human bone marrow protein, SEQ ID NO: 414.
XX KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX OS Homo sapiens.
XX PN WO200153453-A2.
XX PD 26-JUL-2001.
XX PF 23-DEC-2000; 2000WO-US34960.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.

```

```

PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX WPI: 2001-488707/53.
XX DR N-PSDB; AAH90057.
XX PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX for treating e.g. cancer and immune deficiency disorders -
XX PS Claim 10; Page 504-505; 648pp; English.
XX CC The present sequence is one of 251 novel human polypeptides encoded
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX SQ Sequence 366 AA;

Query Match 69.9%; Score 351; DB 22; Length 366;
Best Local Similarity 67.7%; Pred. No. 1e-35;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLQPDPAARYNNLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
Db ||||| ||| : : || : ||||| ||||| ||||| ||||| ||||| : |||
101 mqslspdkaqytsiygalkkimrtgfwprlgvnmimgagpahamyfacyenmkrtl 160

QY 61 SDVIHPGGNSHTANGAAGCVATILLHDAANPAE 93
Db : ||| ||||| ||| || : ||||| |||||
161 ndvfhhqgnshlangiagsmatilhdavnmpae 193

RESULT 9
AAU29748
ID AAU29748 standard; Protein; 677 AA.
XX AC AAU29748;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #239.
XX KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-US08656.
XX PR 18-APR-2000; 2000US-0552929.

```

```
PR 26-JAN-2001; 2001US-0770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy -
XX Claim 20: Page 189; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising
XX the nucleic acids encoding the polypeptides and cells genetically
XX engineered to express them are also useful for producing the proteins.
XX The proteins are useful in genetic vaccination, testing and
XX therapy, and can be used as nutritional supplements. They may be used to
XX increase stem cell proliferation; to regulate haematopoiesis; and in
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX immune suppression and/or stimulation; as anti-inflammatory agents; and
XX in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
XX sequences of novel human secreted proteins of the invention.
XX Sequence 677 AA;
SQ
Query Match 64.1%; Score 322; DB 21; Length 677;
Best Local Similarity 63.4%; Pred. No. 1.1e-21; Indels 0; Gaps 0;
Matches 59; Conservative 13; Mismatches 21;
QY 1 MQSLQPDPAARYNRVLEALWRIIRTEGLWRPMDGLNTATGAGPAHALYFACYEKLKRTL 60
Db 85 mqslspdqspvsiygalckimktgfwrlrgvnmimgagpahamyfacyenmkrtl 144
QY 61 SDVTHPGGNSHIANGAAGCVATLHDDAMNPAAE 93
Db 145 ndvfhggnshlangiagsmatlhhdavmpae 177
RESULT 10
AAB42966
ID AAB42966 standard; Protein: 155 AA.
XX
XX AAB42966;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF2730 polypeptide sequence SEQ ID NO:5460.
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive.
XX Homo sapiens.
OS
PR 26-JAN-2001; 2001US-0770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy -
XX Claim 20: Page 189; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising
XX the nucleic acids encoding the polypeptides and cells genetically
XX engineered to express them are also useful for producing the proteins.
XX The proteins are useful in genetic vaccination, testing and
XX therapy, and can be used as nutritional supplements. They may be used to
XX increase stem cell proliferation; to regulate haematopoiesis; and in
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX immune suppression and/or stimulation; as anti-inflammatory agents; and
XX in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
XX sequences of novel human secreted proteins of the invention.
XX Sequence 677 AA;
SQ
Query Match 64.1%; Score 322; DB 21; Length 677;
Best Local Similarity 63.4%; Pred. No. 1.1e-21; Indels 0; Gaps 0;
Matches 59; Conservative 13; Mismatches 21;
QY 1 MQSLQPDPAARYNRVLEALWRIIRTEGLWRPMDGLNTATGAGPAHALYFACYEKLKRTL 60
Db 85 mqslspdqspvsiygalckimktgfwrlrgvnmimgagpahamyfacyenmkrtl 144
QY 61 SDVTHPGGNSHIANGAAGCVATLHDDAMNPAAE 93
Db 145 ndvfhggnshlangiagsmatlhhdavmpae 177
RESULT 10
AAB42966
ID AAB42966 standard; Protein: 155 AA.
XX
XX AAB42966;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF2730 polypeptide sequence SEQ ID NO:5460.
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive.
XX Homo sapiens.
OS
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX N-PSDB; AAC77175.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX Claim 11; Page 4641; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX Sequence 155 AA;
SQ
Query Match 55.0%; Score 276; DB 21; Length 155;
Best Local Similarity 64.0%; Pred. No. 1.1e-26;
Matches 48; Conservative 13; Mismatches 14; Indels 0; Gaps 0;
QY 1 MQSLQPDPAARYNRVLEALWRIIRTEGLWRPMDGLNTATGAGPAHALYFACYEKLKRTL 60
Db 73 mqslspdqspvsiygalckimktgfwrlrgvnmimgagpahamyfacyenmkrtl 132
QY 61 SDVTHPGGNSHIANG 75
Db 133 ndvfhggnshlang 147
RESULT 11
AAB61130
ID ABB61130 standard; Protein: 379 AA.
XX
XX ABB61130;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 10182.
XX
```


KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
XX	23-MAR-2001; 2001WO-US09231.
Pf	
PF	Venter JC, Adams M, Li PWD, Myers EW;
XX	
XX	WPI; 2001-656860/75.
DR	N-PSDB; ABL05233.
DR	
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Disclosure; SEQ ID NO 10182; 2lpp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 379 AA;
	Query Match 34.2%; Score 171.5; DB 22; Length 379;
	Best Local Similarity 40.9%; Pred. No. 5.2e-13;
	Matches 38; Conservative 16; Mismatches 34; Indels 5; Gaps 2;
QY	1 MQSLQPDAARYNVLEALWRIIRTEGLWRPMRGILNVTATGAGPAHALYFACYEKLKTL 60
	: : : : : : : : : : : : : : : : :
Db	42 mqsl--spptkmlvstmitregllrpigasavvlgagpahsiyfaayemtkelt 99
	: : : : : : : : : : : : : : : : : : :
QY	61 SDVHPGNSHIANGACGCATLLHDAAMNPAE 93
	: : : : : : : : : : : : : : : : : : :
Db	100 akfsvrnlvyisga---vatlhdaissptd 129
RESULT 12	
AAM85891	ID
XX	AAM85891 standard; Protein; 59 AA.
XX	
AC	AAM85891;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen SEQ ID NO:13484.
XX	
KW	Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW	cystostatic; gene therapy; vaccine; metastasis.
XX	
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
XX	
PD	09-AUG-2001.

PR	29-SEP-2000;	2000US-02363659;
PR	29-SEP-2000;	2000US-02363700;
PR	02-OCT-2000;	2000US-02368002;
PR	02-OCT-2000;	2000US-02370337;
PR	02-OCT-2000;	2000US-02370338;
PR	02-OCT-2000;	2000US-02370339;
PR	02-OCT-2000;	2000US-02370400;
PR	13-OCT-2000;	2000US-02399935;
PR	13-OCT-2000;	2000US-02399937;
PR	20-OCT-2000;	2000US-02409660;
PR	20-OCT-2000;	2000US-02412221;
PR	20-OCT-2000;	2000US-02417851;
PR	20-OCT-2000;	2000US-02417861;
PR	20-OCT-2000;	2000US-02417871;
PR	20-OCT-2000;	2000US-02418008;
PR	20-OCT-2000;	2000US-02418009;
PR	20-OCT-2000;	2000US-02418266;
PR	01-NOV-2000;	2000US-02446171;
PR	08-NOV-2000;	2000US-02466474;
PR	08-NOV-2000;	2000US-02466475;
PR	08-NOV-2000;	2000US-02466478;
PR	08-NOV-2000;	2000US-02466477;
PR	08-NOV-2000;	2000US-02466478;
PR	08-NOV-2000;	2000US-02465233;
PR	08-NOV-2000;	2000US-02465234;
PR	08-NOV-2000;	2000US-02465235;
PR	08-NOV-2000;	2000US-02465236;
PR	08-NOV-2000;	2000US-02465237;
PR	08-NOV-2000;	2000US-02465238;
PR	08-NOV-2000;	2000US-02465239;
PR	08-NOV-2000;	2000US-02466002;
PR	08-NOV-2000;	2000US-02466003;
PR	08-NOV-2000;	2000US-02466111;
PR	08-NOV-2000;	2000US-02466133;
PR	17-NOV-2000;	2000US-02492007;
PR	17-NOV-2000;	2000US-02492008;
PR	17-NOV-2000;	2000US-02492009;
PR	17-NOV-2000;	2000US-02492010;
PR	17-NOV-2000;	2000US-02492111;
PR	17-NOV-2000;	2000US-02492112;
PR	17-NOV-2000;	2000US-02492113;
PR	17-NOV-2000;	2000US-02492114;
PR	17-NOV-2000;	2000US-02492115;
PR	17-NOV-2000;	2000US-02492116;
PR	17-NOV-2000;	2000US-02492117;
PR	17-NOV-2000;	2000US-02492937;
PR	17-NOV-2000;	2000US-02492939;
PR	01-DEC-2000;	2000US-02493000;
PR	01-DEC-2000;	2000US-02501660;
PR	01-DEC-2000;	2000US-02503031;
PR	05-DEC-2000;	2000US-02510530;
PR	05-DEC-2000;	2000US-02519888;
PR	06-DEC-2000;	2000US-02567119;
PR	08-DEC-2000;	2000US-02518566;
PR	08-DEC-2000;	2000US-02518568;
PR	08-DEC-2000;	2000US-02518659;
PR	08-DEC-2000;	2000US-02518659;
PR	08-DEC-2000;	2000US-02519890;
PR	11-DEC-2000;	2000US-02519900;
PR	05-JAN-2001;	2000US-02540978;
PR	05-JAN-2001;	2000US-02549677;

	Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -
Claim 11;	SEQ ID NO 13484; 3071pp + Sequence Listing; English.
AAK54951 to AAK64702	encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention.
Sequence	59 AA:
Query Match	33.0%; Score 165.5; DB 22; Length 59;
Best Local Similarity	59.3%; Pred. No. 2.8e-13;
Matches	35; Conservative 1; Mismatches 12; Indels 11; Gaps 1;
QY	8 PAARYRNVLALWRIIRTEGLWRPGRGLNVTATGAGPAHALYFACYEKLKKTLSDVIHP 66 : : Db 12 pldyk-----ggpmearglnvtatgagpxhglyfacteklkalsdvihp 59
RESULT 13	
AAW17054	
ID AAW17054 standard; Protein; 289 AA.	
XX AAW17054;	
DT 05-AUG-1997 (first entry)	
DE Blackcurrant RIB7 polypeptide.	
KW Blackcurrant; fruit-specific promoter; RIB7; transgenic plant.	
OS Ribes nigrum strain Ben Alder.	
XX WO9717452-A1.	
PV 15-MAY-1997.	
FF 04-NOV-1996; 96WO-EP04807.	
XX 03-NOV-1995; 95GB-0022558.	
PP (SMIK) SMITHKLINE BEECHAM PLC.	
PI Brennan RM, Taylor MA, Woodhead MR;	
DR WPI; 1997-281041/25.	
NR N-PSTB; AAT68957.	
XX New isolated promoters from blackcurrant fruit - used for driving fruit-specific expression of DNA sequences in transgenic blackcurrant and other non-climacteric fruit	
PS Example 4; Page 35-36; 66pp; English.	
CC RIB polypeptides (AAW17050-54) are encoded by cDNA clones (AAT68953-	

```
CC 57) that exhibit differential expression in blackcurrant fruit
CC during the ripening period of fruit development. RIB7 shows
CC 62% similarity to yeast MRS4, a yeast mitochondrial RNA splicing
CC protein. RIB7 is expressed almost entirely in fruit. The
CC promoter region (AAT68952) of the RIB7 gene can be used as a
CC fruit-specific promoter.
XX
SQ Sequence 289 AA;

Query Match 26.0%; Score 130.5; DB 18; Length 289;
Best Local Similarity 35.3%; Pred. No. 5.5e-08;
Matches 30; Conservative 17; Mismatches 35; Indels 3; Gaps 1;

QY 9 AARYRNVLALMRIITEGLWRPMRGLNVTTATGAGPAHALYFACYEKLKTKTSLSDVTHPGG 68
   :|: :||| :|: ||| :|: |||||:|: ||| :|: |
Db 32 saqsaglrqalsilkevpaglrylganglgagpahavvyfsvyemcketfs---hgdp 88

QY 69 NSHIANGAAGCVATLLHDDAAMNPAE 93
   :|: :|: ||| :|: ||| :|:
Db 89 snsgahavsgvfatvasdavitpmd 113

RESULT 14
AAG22079
ID AAG22079 standard; Protein; 278 AA.
XX
AC AAG22079;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24871.
XX
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
```

```
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159233.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 25.9%; Score 130; DB 21; Length 278;
Best Local Similarity 37.9%; Pred. No. 6e-08;
Matches 36; Conservative 15; Mismatches 36; Indels 8; Gaps 3;

Oy 1 MOSLPDPAARYNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
    |||: | : : : |||:| ||: |||||:|: || | |
Db 13 mqaIrcp-ikpigrqafsrisktdgpsalyrgiwamglgagpahavyfsfyevskkf 71
    |||: | : : : |||:| ||: |||||:|: || | |

Oy 61 SDVIHPGG--NSHIANGAAGCVATLLHDAANPAE 93
    | || | : | : | | | : | | | :
Db 72 s-----ggnpnnsaahaisgvfatissdavftpm 101

RESULT 15
AAG43094
ID AAG43094 standard; Protein; 278 AA.
XX
AC AAG43094;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53822.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
```

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 25.9%; Score 130; DB 21; Length 278;
Best Local Similarity 37.9%; Pred. No. 6e-08;
Matches 36; Conservative 15; Mismatches 36; Indels 8; Gaps 3;

Qy 1 MQLQPDPAARYRNVLALWRIIRTEGLWRPMPGILNVTATCAGPAHALYFACYEKLKKTLL 60
|||: | : : |||: | : |||::|: | | |
Db 13 mqlrscp-ikpigirqafsrliktdgpsalrygiwamglgagpahavfysfyevskfl 71

QY 61 SDVIHPGG--NSHIANGAAGCVATLLHDAAMNPAE 93
| | | | | : | | : | | : | | :
Db 72 s-----gggnpnnsaahisgvfatissdvaftpm 101

Search completed: August 27, 2002, 03:56:00
Job time: 4912 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:58:51 ; Search time 48.11 seconds
(without alignments)
47.724 Million cell updates/sec

Title: US-09-870-113-10
Perfect score: 502
Sequence: 1 MQLQPPDPAARYRNVLALW.....GAAGCVATLLHDAAMPAEG 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	130.5	26.0	289	4 US-09-068-140A-10	Sequence 10, Appl
2	130.5	26.0	328	4 US-09-068-140A-15	Sequence 15, Appl
3	89.5	17.8	351	2 US-08-933-750C-19	Sequence 19, Appl
4	89.5	17.8	351	4 US-09-234-613-19	Sequence 19, Appl
5	89	17.7	320	2 US-08-933-750C-12	Sequence 12, Appl
6	89	17.7	320	4 US-09-234-613-12	Sequence 12, Appl
7	71	14.1	2584	3 US-08-936-135-4	Sequence 4, Appl
8	69	13.7	306	5 PCT-US94-09799-1	Sequence 1, Appl
9	68.5	13.6	312	4 US-09-142-565-2	Sequence 2, Appl
10	67	13.3	432	2 US-08-937-466-4	Sequence 4, Appl
11	67	13.3	432	2 US-09-172-528-4	Sequence 4, Appl
12	67	13.3	432	3 US-09-318-198-4	Sequence 4, Appl
13	67	13.3	432	4 US-09-503-573-4	Sequence 4, Appl
14	65.5	13.0	311	2 US-08-775-009-33	Sequence 33, Appl
15	65	12.9	1094	4 US-09-268-347-32	Sequence 32, Appl
16	63	12.5	308	2 US-08-937-466-2	Sequence 2, Appl
17	63	12.5	308	2 US-09-172-528-2	Sequence 2, Appl
18	63	12.5	308	3 US-09-318-199-2	Sequence 2, Appl
19	63	12.5	308	4 US-09-503-573-2	Sequence 2, Appl
20	63	12.5	311	2 US-08-775-009-32	Sequence 32, Appl
21	63	12.5	447	4 US-09-160-119-4	Sequence 4, Appl
22	63	12.5	571	2 US-08-892-770-5	Sequence 5, Appl
23	63	12.5	674	4 US-09-160-119-2	Sequence 2, Appl
24	62	12.4	469	4 US-09-188-930-339	Sequence 339, App
25	61.5	12.3	2588	3 US-08-936-135-2	Sequence 2, Appl
26	61	12.2	517	1 US-08-190-802A-30	Sequence 30, Appl
27	61	12.2	517	4 US-08-477-346-30	Sequence 30, Appl

28 61 12.2 517 4 US-08-473-089-30 Sequence 30, Appl
29 60.5 12.1 256 2 US-08-937-466-6 Sequence 6, Appl
30 60.5 12.1 256 2 US-09-172-528-6 Sequence 6, Appl
31 60.5 12.1 256 3 US-09-318-199-6 Sequence 6, Appl
32 60.5 12.1 256 4 US-09-503-579-6 Sequence 6, Appl
33 60 12.0 1095 4 US-09-112-096-15 Sequence 15, Appl
34 59 11.8 133 1 US-07-917-722-3 Sequence 3, Appl
35 59 11.8 565 2 US-08-892-770-6 Sequence 6, Appl
36 58 11.6 236 3 US-09-248-335-68 Sequence 68, Appl
37 58 11.6 299 1 US-08-518-878B-56 Sequence 56, Appl
38 58 11.6 299 2 US-08-470-868A-56 Sequence 56, Appl
39 58 11.6 309 1 US-08-518-878B-51 Sequence 51, Appl
40 58 11.6 309 2 US-08-807-861A-51 Sequence 51, Appl
41 58 11.6 309 2 US-08-470-868A-51 Sequence 51, Appl
42 58 11.6 309 3 US-09-210-681-51 Sequence 51, Appl
43 58 11.6 309 3 US-08-946-719A-51 Sequence 51, Appl
44 58 11.6 312 4 US-09-188-930-142 Sequence 142, App
45 58 11.6 372 1 US-08-278-729A-33 Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-068-140A-10
; Sequence 10, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-10

```
Query Match          26.0%; Score 130.5; DB 4; Length 289;
Best Local Similarity 35.3%; Pred. No. 1.5e-08;
Matches 30; Conservative 17; Mismatches 35; Indels 3; Gaps 1;

Qy 9 AARYRNVLALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKLTLSVDVHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 SAQSAGLRQALGSILKVEGPAGLYRGIGAMGLGAGPAHAHVFSVYEMCKETFS---HGDP 88

Qy 69 NSHIANGAAGCVATLLHDAAMNPAE 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 SNSGAHAVSGVFATVSDAVITPMD 113

RESULT 2
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068.140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-15

Query Match          26.0%; Score 130.5; DB 4; Length 328;
Best Local Similarity 35.3%; Pred. No. 1.8e-08;
Matches 30; Conservative 17; Mismatches 35; Indels 3; Gaps 1;

Qy 9 AARYRNVLALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKLTLSVDVHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 SAQSAGLRQALGSILKVEGPAGLYRGIGAMGLGAGPAHAHVFSVYEMCKETFS---HGDP 88

Qy 69 NSHIANGAAGCVATLLHDAAMNPAE 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 SNSGAHAVSGVFATVSDAVITPMD 113

RESULT 2
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068.140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-15

Query Match          26.0%; Score 130.5; DB 4; Length 328;
Best Local Similarity 35.3%; Pred. No. 1.8e-08;
Matches 30; Conservative 17; Mismatches 35; Indels 3; Gaps 1;

Qy 9 AARYRNVLALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKLTLSVDVHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 SAQSAGLRQALGSILKVEGPAGLYRGIGAMGLGAGPAHAHVFSVYEMCKETFS---HGDP 88

Qy 69 NSHIANGAAGCVATLLHDAAMNPAE 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 SNSGAHAVSGVFATVSDAVITPMD 113

RESULT 3
US-08-933-750C-19
; Sequence 19, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933.750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNOAT01
; CLONE: 724157
US-08-933-750C-19

Query Match          17.8%; Score 89.5; DB 2; Length 351;
Best Local Similarity 32.3%; Pred. No. 0.0029;
Matches 31; Conservative 12; Mismatches 32; Indels 21; Gaps 5;

Qy 7 DPAARYRNVLALWRIIRTEG---LWRPRLGNVTATGAGPAHALYFACYEKLKLT--- 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 DP-TRFTGTMDFAFKIVRHEGRTLW---SGLPATLVMTVPATAIYFTAYDQLKAFLCGR 147

Qy 61 ---SDVTHPGNSHIANGAAGCVATLLHDAAMNPAE 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 ALTSDLYAP-----MWAGALARLGTVTVISPLE 175
```



```
RESULT 4
US-09-234-613-19
; Sequence 19, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYN00AT01
; CLONE: 724157
US-09-234-613-19

Query Match 17.8%; Score 89.5; DB 4; Length 351;
Best Local Similarity 32.3%; Pred. No. 0.0029;
Matches 31; Conservative 12; Mismatches 32; Indels 21; Gaps 5;

QY 7 DPAARYRNVLALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLLKTL--- 60
DB 92 DP-TRETGTMDAFVKIVRHGHTLW---SGLPATLMTVTPATAIYFTAYDQLKAFLCGR 147

QY 61 ---SDVHPGGNSHANGAGCVATLLHDAAMNPAE 93
DB 148 ALTSDLYAP-----MWAGALARLGTVTISPLE 175

RESULT 5
US-08-933-750C-12
; Sequence 12, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
```

```
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNOT02
; CLONE: 207452
US-08-933-750C-12

Query Match 17.7%; Score 89; DB 2; Length 320;
Best Local Similarity 27.3%; Pred. No. 0.0029;
Matches 24; Conservative 18; Mismatches 28; Indels 18; Gaps 4;

QY 7 DPAARYRNVLALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLLKTLSDV 63
DB 54 DSAKYHGILOASRQILQEGPTAFKGVHPAQILSIGYGAVOFLSF-----EMLTEL 106

QY 64 IHPGG-----NSH-IANGAGCVATL 83
DB 107 VHRGSVYDAREFSVHFVCGGLAACMATL 134

RESULT 6
US-09-234-613-12
; Sequence 12, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
```

APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/234,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SPLNNOT02
CLONE: 207452
US-09-234-613-12

Query Match 17.7%; Score 89; DB 4; Length 320;

Best Local Similarity 27.3%; Pred. No. 0.0029;

Matches 24; Conservative 18; Mismatches 28; Indels 18; Gaps 4;

QY 7 DPAAARNVLEALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKKTLSDV 63

DB 54 DPSAKYHGLQASRQIQLEEGTAFMKGHVPAQILSIGAVQFLSF-----EMLTEL 106

QY 64 IHPGG-----NSH-IANGAAGCVATL 83

DB 107 VHRGSVDAREFSVHFVCGGLAACMATL 134

RESULT 7

US-08-936-135-4

; Sequence 4, Application US/089361135

; Patent No. 6054293

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: He, Zhigang

; APPLICANT: Chen, Hang

; TITLE OF INVENTION: Semaphorin Receptors

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC97-288-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 343-4341

; TELEFAX: (650) 343-4342

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2584 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-936-135-4

Query Match 14.1%; Score 71; DB 3; Length 2584;

Best Local Similarity 29.1%; Pred. No. 8.2;

Matches 32; Conservative 15; Mismatches 35; Indels 28; Gaps 6;

QY 3 SLQDDP-AARYNRVLEALWRIIR-TEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60

DB 1320 ALEPRSRPRHSPTRYRILEASNGLRPLEGLNVAL--ASPLEGLY-ASPGLYSL 1376

QY 61 SDVHPGG-----NSHIANGAAGCVATLLHDAAM 89

DB 1377 EVALARGGLYVALILEILEGLNGLYLYSHIS--ARGGLASNLYSVAL 1424

RESULT 8

PCT-US94-09799-1

; Sequence 1, Application PC/TUS9409799

; GENERAL INFORMATION:

; APPLICANT: Kagan, David

; TITLE OF INVENTION: Method and Composition for Weight Reduction

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Watov & Kipnes, P.C.

; STREET: 186 Princeton-Hightstown Rd, PO Box 247

; CITY: Princeton Junction

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 08550

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: DOS 6.0

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/09799

; FILING DATE: 29-AUG-1994

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/249,179

; FILING DATE: 25 May 1994

; APPLICATION NUMBER: 08/114,313

; FILING DATE: 30 August 1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Kipnes, Allen R.

; REGISTRATION NUMBER: 28,433

; REFERENCE/DOCKET NUMBER: 489.1.001A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 609-243-0330

```

1 / APPLICANT: Amaral, M. Catherine
2 / APPLICANT: Chen, Jin-Long
3 / TITLE OF INVENTION: UCP3 Genes
4 / NUMBER OF SEQUENCES: 6
5 / CORRESPONDENCE ADDRESS:
6 / ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
7 / STREET: 75 DENISE DRIVE
8 / CITY: HILLSBOROUGH
9 / STATE: CALIFORNIA
10 / COUNTRY: USA
11 / ZIP: 94010
12 /
13 / COMPUTER READABLE FORM:
14 / MEDIUM TYPE: Floppy disk
15 / COMPUTER: IBM PC compatible
16 / OPERATING SYSTEM: PC-DOS/MS-DOS
17 / SOFTWARE: PatentIn Release #1.0, Version #1.30
18 / CURRENT APPLICATION DATA:
19 / APPLICATION NUMBER: US/08/937,466
20 /
21 / FILING DATE:
22 / CLASSIFICATION: 435
23 / ATTORNEY/AGENT INFORMATION:
24 / NAME: OSMAN, RICHARD A
25 / REGISTRATION NUMBER: 36,627
26 / REFERENCE/DOCKET NUMBER: T97-009
27 / TELECOMMUNICATION INFORMATION:
28 / TELEPHONE: (650) 343-4341
29 / TELEFAX: (650) 343-4342
30 / INFORMATION FOR SEQ ID NO: 4:
31 / SEQUENCE CHARACTERISTICS:
32 / LENGTH: 432 amino acids
33 / TYPE: amino acid
34 / STRANDEDNESS: single
35 / TOPOLOGY: linear
36 / MOLECULE TYPE: peptide
37 /
38 / US-08-937-466-4
39 /
40 /
41 / Query Match 13.3%; Score 67; DB 2; Len
42 / Best Local Similarity 28.6%; Pred No. 2.6;
43 / Matches 16; Conservative 11; Mismatches 29;
44 /
45 /
46 / QY 8 PAARYRNVLALWRIIRTEGLMRPRLGNLVNTATGAGPAHAYFA
47 / | ||| : | : : || : | : | : | : | : |
48 / DB 245 PLGRYSPLHCLMKVAQSGPTAFYKGFVPSFLRLGAWNMVMTF
49 /
50 /
51 / RESULT 11
52 / US-09-172-528-4
53 / Sequence 4, Application US/09172528
54 / Patent No. 5952469
55 / GENERAL INFORMATION:
56 / APPLICANT: Zhang, Ning
57 / APPLICANT: Amaral, M. Catherine
58 / APPLICANT: Chen, Jin-Long
59 / TITLE OF INVENTION: UCP3 Genes
60 / NUMBER OF SEQUENCES: 6
61 / CORRESPONDENCE ADDRESS:
62 / ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
63 / STREET: 75 DENISE DRIVE
64 / CITY: HILLSBOROUGH
65 / STATE: CALIFORNIA
66 / COUNTRY: USA
67 / ZIP: 94010
68 /
69 / COMPUTER READABLE FORM:
70 / MEDIUM TYPE: Floppy disk
71 / COMPUTER: IBM PC compatible
72 / OPERATING SYSTEM: PC-DOS/MS-DOS
73 / SOFTWARE: PatentIn Release #1.0, Version #1.30
74 / CURRENT APPLICATION DATA:
75 / APPLICATION NUMBER: US/09/172,528
76 / FILING DATE:
77 / CLASSIFICATION:
78 / PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-172-528-4

Query Match 13.3%; Score 67; DB 2; Length 432;
Best Local Similarity 28.6%; Pred.No. 2.6;
Matches 16; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

QY 8 PAARYNVLEALWRIIRTEGLRPMRLGNVTATGAGPAHALYFACYEYKLLKKTLSDV 63
DB 245 PLGYRSPFLCHMLKMWAGEGTATYKGFVPSFLRLGAWNVMFVTEYQLKRALMKV 300
| ||| : : : : || : : : : | : : : : || : || : |
| : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 12
US-09-318-199-4
; Sequence 4, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/318,199
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-318-199-4

```

APPLICANT: Roe, Bruce
TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
TITLE OF INVENTION: Veliocardiofacial Syndrome Minimal Critical Region
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSER: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSER: No. 5935783ris, LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,009
FILING DATE: 27-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CH-0681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-009-33

Query Match 13.0%; Score 65.5; DB 2; Length 311;
Best Local Similarity 30.6%; Pred. No. 2.6;
Matches 19; Conservative 12; Mismatches 28; Indels 3; Gaps 1;
QY 4 LQDPDPA---RYRNLALWRIIRTEGLWRPWRGLNVTATGAGPAHALYFACYEKLKLT 60
DB 53 LQIDERANPPRYRGIDCVQRQVRSVGLYRGLSSLLYGSIPRAAVRFGMFEFLSNEM 112
QY 61 SD 62
DB 113 RD 114

RESULT 15
US-09-268-347-32
Sequence 32, Application US/09268347
Patent No. 6335182
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
FILE REFERENCE: 1038-860
CURRENT APPLICATION NUMBER: US/09/268,347
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32
LENGTH: 1094
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-268-347-32

Query Match 12.9%; Score 65; DB 4; Length 1094;
Best Local Similarity 24.7%; Pred. No. 15;
Matches 22; Conservative 12; Mismatches 23; Indels 32; Gaps 3;

QY 34 GLNVTATGAGPAHALYFACYEKL---KKTLSVDVIHPGNS-----H 71
DB 150 GVOVTSTSENGKHAIITFALAKDLDMTATVSDTLTIGSGSATPKVNVNSTASGLN 209
QY 72 IANGAAGC-----VATLIHDAAMN 90
DB 210 FAKGATGANGDTTVVHLTNIASTLQDTLLN 238
Search completed: August 27, 2002, 03:58:53
Job time: 4895 sec

A;Accession: A86205
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-781
A;Cross-references:
C;Genetics:
A;Map position: 1

```
Query Match      26.9%; Score 135; DB 2; Length 781;
Best Local Similarity 38.9%; Pred. No. 6.8e-07;
Matches 37; Conservative 14; Mismatches 36; Indels 8; Gaps 3;

1 MOSLOPDAARYNVYLEALWRIIRTEGWLWRPMRGUNVTATGAGPAHAIFYACYEKLKRTL 60
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
64 MOALRPCP-LKPVGIREFRSTIQEGPSALYIGIWMGLGAGPAHAIFYFSFEYSKKYL 122
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

61 SDVIHPGG--NSHIANGAACGVATLLDHAAMNPAE 93
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

123 S-----AGDONNSVAHAMSGVFATISSDAVTFPMD 152
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```

RESULT

T00582

probable mitochondrial carrier protein [unpublished] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T27E13.10
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
C:Accession: T00582; C84705
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC T27E13 genomic sequence.
A:Reference number: Z14178
A:Accession: T00582
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-331 <FOU>
A:Cross-references: EMBL:AC004165; NID:g3150396; PID:g3150404
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20093487
A:Accession: C84705
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <STO>
A:Cross-references: GB:AE002093; NID:g3150404; PIDN:AAC16956.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g30160; T27E13.10
A:Map position: 2
A:Introns: 263/3
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: mitochondrion

Query Match 25.9%; Score 130; DB 2; Length 331;
Best Local Similarity 37.9%; Pred. No. 9.7e-07;
Matches 36; Conservative 15; Mismatches 36; Indels 8; Gaps 3;

Qy 1 MOSLPDPAARYNVLEALWRIIRTEGLWRPARGLNVTATGAGPAHALFPACYEKIKKTL 60
|||: | : : ||::| |||::|: | | |
Dd 66 MOALRSCP- IKPICROAFSIIKTDPGSALYRGIVANGLGAGPAHAVFSFYEVSKFEL 124

QY 61 SDVIHPGG--NSHIANGAAGCVATILHDAAMNPAE 93
| | | | | : | | | | : | | :
Db 125 S-----GPNPNSAAHAISGVFATILSSDAVFETPMD 151

RESULT
S55179

mtRNA splice defect-suppressing mitochondrial carrier MRS3 - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: protein J0675; protein YJL133W

C;Species: *Saccharomyces cerevisiae*

C;Date: 08-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 21-Jul-2000

C;Accession: S55179; S01267; S20228; S56915; S71664

R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.

submitted to the EMBL Data Library, May 1995

yeast hypothetical proteins.

A;Reference number: S55159

A;Accession: S55179

A;Molecule type: DNA

A;Residues: 1-314 <KAT>

A;Cross-references: EMBL:X87371; NID:g854542; PID:g854563

R;Schmidt, C.; Soellner, T.; Schweyen, R.J.

Mol. Gen. Genet. 210, 145-152, 1987

A;Title: Nuclear suppression of a mitochondrial RNA splice defect: nucleotide sequence

A;Reference number: S01267; MUID:88121698

A;Accession: S01267

A;Molecule type: DNA

A;Residues: 1-86, 'S', 88-102, 'L', 104-127, 'Y', 129-141, 'M', 143-187, 'R', 189-314 <SCH>

A;Cross-references: EMBL:X06239; NID:g3990; PIDN:CAA29582.1; PID:g3991

R;Wiesenberg, G.; Link, T.A.; von Ahsen, U.; Waldherr, M.; Schweyen, R.J.

J. Mol. Biol. 217, 23-37, 1991

A;Title: MRS3 and MRS4, two suppressors of mtRNA splicing defects in yeast, are new m

A;Reference number: S13532; MUID:91108815

A;Accession: S20228

A;Molecule type: DNA

A;Residues: 15-314 <WIE>

A;Cross-references: EMBL:X56445; NID:g3992; PIDN:CAA39830.1; PID:g3994

R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S56912

A;Accession: S56915

A;Molecule type: DNA

A;Residues: 1-314 <RAW>

A;Cross-references: EMBL:Z49408; NID:g1008337; PID:g1008338; MIPS:YJL133w

R;Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.

Yeast 12, 787-797, 1996

A;Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome XI.

A;Reference number: S71643; MUID:96408771

A;Accession: S71664

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-314 <KAF>

A;Cross-references: EMBL:X87371; NID:g854542; PIDN:CAA60822.1; PID:g854563

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995

C;Genetics:

A;Gene: SGD:MRS3

A;Cross-references: SGD:S0003669; MIPS:YJL133w

A;Map position: 10L

A;Genome: nuclear

C;Function:

C;Description: probably involved in splicing of all intron for COB gene; essential for

licing

C;Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology

C;Keywords: mitochondrial inner membrane; mitochondrion; pre-mRNA splicing

F;30-119/Domain: ADP, ATP carrier protein repeat homology <ACP1>

F;127-211/Domain: ADP, ATP carrier protein repeat homology <ACP2>

F;216-311/Domain: ADP, ATP carrier protein repeat homology <ACP3>

Query Match	24.7%	Score 124;	DB 2;	Length 314;
Best Local Similarity	36.7%;	Pred. No. 4.3e-06;		
Matches	29;	Conservative	9;	Mismatches 41; Indels 0; Gaps
Qy	13	RNVLEALWRIIRTEGLMRPRLGNTVATGAGPAHALYFACYEKLKKTLSVDVIHPGGNSHI	72	
Db	72	KNMLLSQISHISTSGTLLALWKQVSVILGAGPAHAVYFGTYECKKMLIDSSDTQTHHPF	131	
Qy	73	ANGAAGCVCATILLHDAMNP	91	

Db 132 KTAISGACATTASDALMNP 150

RESULT 5
T39149
probable RNA splicing proteinmitochondrial carrier protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Accession: T39149
A:Reference number: Z21748
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-303 <OLI>
A:Cross-references: EMBL:599168; PIDN:CAB16300.1; GSPDB:GN00066; SPDB:SPAC8C9.12c
A:Experimental source: strain 972h-; cosmid c8C9
C:Genetics:
A:Map position: 1
A:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 23.5%; Score 118; DB 2; Length 303;
Best Local Similarity 33.3%; Pred. No. 1.9e-05;
Matches 26; Conservative 17; Mismatches 31; Indels 4; Gaps 1;

QY 14 NVLEALWRIIRTEGLWRPRLGNTATGAGPAHALYFACYEKLKLTLSVDVIHPGNSHIA 73
DB 60 NIVNSVIKISSTEGVYSLWRGISSVIMGAPSHAIYFVLEFFKSK----INASPRPLA 115
QY 74 NGAGCVATLLHDAAMNP 91
DB 116 SALAGACAITISDAFMT 133

RESULT 6
S13533
mRNA splice defect-suppressing mitochondrial carrier MRS4 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKR052c
C:Species: Saccharomyces cerevisiae
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jul-2000
R:Wiesenberg, G.; Link, T.A.; von Ahsen, U.; Waldherr, M.; Schweyen, R.J.
J. Mol. Biol. 217, 23-37, 1991
A:Title: MRS3 and MRS4, two suppressors of mRNA splicing defects in yeast, are new members of the ADP/ATP carrier family
A:Reference number: S13533; MUID:91108815
A:Accession: S13533
A:Molecule type: DNA
A:Residues: 1-304 <JMO>
A:Cross-references: EMBL:X56444; NID:g3995; PIDN:CAA39828.1; PID:g3996
R:Visser, S.; Urrestarazu, L.A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38118
A:Accession: S38126
A:Molecule type: DNA
A:Residues: 1-304 <VIS>
A:Cross-references: EMBL:Z28277; NID:g486506; PID:g486507; MIPS:YKR052c
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:MRS4
A:Cross-references: SGD:S0001760; MIPS:YKR052c
A:Map position: 11R
A:Genome: nuclear
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: duplication; mitochondrion; transmembrane protein
F:206-109/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:117-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:206-301/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 21.1%; Score 106; DB 2; Length 304;

Best Local Similarity 32.5%; Pred. No. 0.00041;
Matches 25; Conservative 12; Mismatches 40; Indels 0; Gaps 0;

QY 15 VLEALWRIIRTEGLWRPRLGNTATGAGPAHALYFACYEKLKLTLSVDVIHPGNSHIAN 74
DB 64 MISTQISKISTMEGSMALWKGVSVILGAGPAHAVYFTGYEFCARLISPEDMOTHPMKT 123
QY 75 GAAGCVATLLHDAAMNP 91
DB 124 AUSGTIATTAADALMNP 140

RESULT 7

S44092
probable carrier protein c2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Oct-1999
A:Accession: S44092; T24753
R:Runswick, M.J.; Philipides, A.; Lauria, G.; Walker, J.E.
submitted to the EMBL Data Library, November 1993
A:Description: Extension of the mitochondrial transport superfamily: sequences of five members of the ADP/ATP carrier family
A:Reference number: S44090
A:Accession: S44092
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <RUN>
A:Cross-references: EMBL:X76116; NID:g472899; PIDN:CAA53722.1; PID:g472900
R:Lloyd, C.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z19932
A:Accession: T24753
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-384 <WIL>
A:Cross-references: EMBL:Z49070; PIDN:CAA88869.1; GSPDB:GN00020; CESP:T09F3.2
A:Experimental source: clone T09F3
C:Genetics:
A:Gene: CESP:T09F3.2
A:Map position: 2
A:Introns: 82/1; 113/1; 153/2; 179/2; 259/3; 359/3
C:Superfamily: Caenorhabditis probable carrier protein c2; ADP,ATP carrier protein repeat homology <ACP1>
F:4-43,141-193/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:45-67/Region: serine-rich
F:201-285/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:296-381/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 20.2%; Score 101.5; DB 1; Length 384;
Best Local Similarity 35.1%; Pred. No. 0.0017;
Matches 33; Conservative 17; Mismatches 31; Indels 13; Gaps 5;

QY 5 QPDPAARYRN-VLEALWRIIRTEGLWRPRLGNTATGAGPAHALYFACYEKLKLTLS 61
DB 137 QPPTAARRGTIVIKYITQVIKTEGIGALYKGLIPNLGVAPSKAVFYTYTSKRFWNES 196
QY 62 DVTHPGNSHIAN----GAAGCVATLLHDAAMNP 91
DB 197 EVLIP--NSAIVHVSAGSAGFVAA----SAVNP 224

RESULT 8

S60949
probable phosphate transport protein, mitochondrial - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein OS064; protein YOR222w; protein YOR50-12
C:Species: Saccharomyces cerevisiae
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 21-Jul-2000
R:Galisson, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A:Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome 1
A:Reference number: S60938
A:Accession: S60949
A:Molecule type: DNA

Query Match 18.8%; Score 94.5; DB 2; Length 296;
Best Local Similarity 37.9%; Pred. No. 0.0077;
Matches 25; Conservative 8; Mismatches 20; Indels 13; Gaps 3;

QY 22 IIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLSDVTHPG-----GNSH----I 72
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150 ILRRQGSLGYRLGITITVLDRDAPAHGLFYFTYEVRERL-----HPGCRKTGQENLRMTLV 205
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 73 ANGAAG 78
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 206 AGGLAG 211

RESULT 10
D84613
hypothetical protein At2g22500 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84613
R:/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:/Reference number: A84420; MUID:20083487
A:/Accession: D84613
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-313 <STO>
A:/Cross-references: GB:AE002093; NID:g544443; PIDN:AAD22351.1; GSPDB:GN00139
C:/Genetics:
A:/Map position: 2
C:/Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology

Query Match 18.8%; Score 94.5; DB 2; Length 313;
Best Local Similarity 31.8%; Pred. No. 0.0081;
Matches 27; Conservative 23; Mismatches 28; Indels 7; Gaps 3;

QY 12 YRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLS--VIHPSGN 69
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 161 YKSVDLAITOMIRGEVTSLWRGSSLTINRAMLVTSQLASYSVKETILEKGLLDGLG 220
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 70 SHI-ANGAACGVATLLHDAANPAE 93
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 221 THVSASFAGFVASV----ASNPD 241

RESULT 11
JQ1459
Bt1 protein precursor - maize
C:Species: Zea mays (maize)
C:/Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 24-Sep-1999
C:/Accession: JQ1459
R:/Sullivan, T.D.; Strelow, L.I.; Illingworth, C.A.; Phillips, R.L.; Nelson Jr., O.E.
Plant Cell 3, 1337-1348, 1991
A:/Title: Analysis of maize brittle-1 alleles and a defective suppressor-mutator-induc
A:/Reference number: JQ1459; MUID:93005685
A:/Accession: JQ1459
A:/Molecule type: mRNA
A:/Residues: 1-436 <SUL>
A:/Cross-references: GB:M79333; NID:g168425; PIDN:AAA33438.1; PID:g168426
C:/Comment: This protein acts as an adenylate translocator in amyloplasts.
C:/Genetics:
A:/Introns: 206/3; 260/3
C:/Superfamily: Bt1 protein; ADP, ATP carrier protein repeat homology
C:/Keywords: chloroplast; duplication; transmembrane protein
F:/1-75/Domain: transit peptide (amyloplast) #status predicted <TNP>
F:/76-436/Product: Bt1 protein #status predicted <MAT>
F:/131-217/Domain: ADP, ATP carrier protein repeat homology <ACPI>

Matches 27; Conservative 14; Mismatches 36; Indels 15; Gaps 3;

Qy 4 LQPDPA-ARYRNVEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKK---- 58

. Db 125 IQDDKAHTKENGPIDATKQLLRTHGLKSLTRGFLATVARDAPAFGVYFASYEWARSCK 184

Qy 59 -----TLSDVIHPGNSHIANGAAGCVATLLH 85

Db 185 DGETSTLS-----SGQLLFAGGTAGMLSWLFN 211

Search completed: August 27, 2002, 03:57:43
Job time: 4949 sec


```

CC membranes of mitochondria. Can transport 2-oxoadipate, 2-
CC oxoglutarate, adipate, glutarate, 2-oxopimelate, oxaloacetate,
CC citrate and malate. The main physiological role is probably to
CC supply 2-oxoadipate and 2-oxoglutarate from the mitochondrial
CC matrix to the cytosol where they are used in the biosynthesis of
CC lysine and glutamate, respectively, and in lysine catabolism.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J75130; CAA99440.1; -.
DR EMBL; X92441; CAA63185.1; -.
DR SGD; S0005748; ODC2.
DR InterPro; IPR002067; Mit_carrier.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
SQ SEQUENCE 307 AA; 34007 MW; 4089082A64DBA97C CRC64;

Query Match 18.9%; Score 95; DB 1; Length 307;
Best Local Similarity 27.7%; Pred. No. 0.0015;
Matches 23; Conservative 17; Mismatches 43; Indels 0; Gaps 0;

QY 11 RYRNVLALWRIIRTEGLWRPGRGLNVTATGAGPAHALYFACYEKLKLTSDVIHPGGNS 70
DB 58 RYGVNIDLUKKIVKGGFSLRGISSPMLMEAPKATKFCNDQYQKIFKNLNTNETT 117

QY 71 HTAAGAAGCVATILHDAANPAE 93
DB 118 QKISIAAGASAGTAAVIVPFE 140

RESULT 4
BT1_MAIZE
ID BT1_MAIZE STANDARD; PRT; 436 AA.
AC P29518;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brittle-1 protein, chloroplast precursor.
GN BT1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005685; PubMed=1668652;
RA Sullivan T.D., Strelow L.I., Ilingworth C.A., Phillips R.L.,
RA Nelson O.E. Jr.;
RT "Analysis of maize brittle-1 alleles and a defective
RT suppressor-mutator-induced mutable allele.";
RL Plant Cell 3:1337-1348(1991).
CC -!- FUNCTION: COULD PLAY A ROLE IN AMYLOPLAST MEMBRANE TRANSPORT.
CC -!- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
CC -!- TISSUE SPECIFICITY: ENDOSPERM OF DEVELOPING KERNELS.

```

```

CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M79333; AAA33438.1; -.
DR PIR; JQ1459; JQ1459.
DR MaizeDB; 47578; -.
DR InterPro; IPR002067; Mit_carrier.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Transit peptide; Chloroplast; Amyloplast; Transmembrane.
FT TRANSIT 1 75 CHLOROPLAST (POTENTIAL).
FT CHAIN 76 436 BRITTLE-1 PROTEIN.
FT TRANSMEM 229 247 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
SQ SEQUENCE 436 AA; 46627 MW; 9600C05F603E9DAE CRC64;

Query Match 18.5%; Score 93; DB 1; Length 436;
Best Local Similarity 41.2%; Pred. No. 0.0038;
Matches 21; Conservative 7; Mismatches 23; Indels 0; Gaps 0;

QY 12 YRNVLALWRIIRTEGLWRPGRGLNVTATGAGPAHALYFACYEKLKLTSD 62
DB 365 YQNVLHAIYCIILKKEGAGGLYRGLGPGSCIKLMPAGIAFCYACKILVD 415

RESULT 5
CG69_HUMAN
ID CG69_HUMAN STANDARD; PRT; 359 AA.
AC Q9BZJ4; Q9UF66; Q9Y379; Q9P182;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial carrier protein CGI-69 (PRO2163).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SUBCELLULAR LOCATION, AND
RP TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX PubMed=11139402;
RA Yu X.X., Lewin D.A., Zhong A., Brush J., Schow P.W., Sherwood S.W.,
RA Pan G., Adams S.H.;
RT "Overexpression of the human 2-oxoglutarate carrier lowers
RT mitochondrial membrane potential in HEK-293 cells; contrast with the
RT unique cold-induced mitochondrial carrier CGI-69.";
RL Biochem. J. 353:369-375(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Cassenhuber J., Glassl S.,
RA Ansoerg W., Boecker M., Bloecker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,

```

RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RL analysis of 500 novel complete protein coding human cDNAs";
RN Genome Res. 11:422-435(2001).
[4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA TISSUE=Brain, and Colon;
RC Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE OF 15-359 FROM N.A. (ISOFORM 1).
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
RA He F.;
RT "Functional prediction of the coding sequences of 79 new genes deduced
by analysis of cDNA clones from human fetal liver.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/CGI-69L (shown here) and 2;
are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Expressed in many tissues. Abundant in testis
and kidney.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF317711; AAG60687.1; -
DR EMBL; AF151827; AAD34064.1; -
DR EMBL; AL133584; CAB63728.1; -
DR EMBL; AK026060; BAB15341.1; -
DR EMBL; BC001398; AAH01398.1; -
DR EMBL; BC009330; AAH09330.1; -
DR EMBL; AF119864; AAF69618.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr. 3.
DR PRINTS; PR00926; MITOCARRIER
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
KW Alternative splicing; Polymorphism.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT VARSPLOC 64 72 LPSSLQSTG -> W (IN ISOFORM 2).
FT VARIANT 247 247 L -> F.
FT FTId=VAR_012756.
FT CONFLICT 266 266 G -> R (IN REF. 6).
SQ SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;

Query Match 17.8%; Score 89.5; DB 1; Length 359;
Best Local Similarity 32.3%; Pred. No. 0.0077;
Matches 31; Conservative 12; Mismatches 32; Indels 21; Gaps 5;
QY 7 DPAARYNRVLEALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL--- 60

Db 100 DP-TRFTGTMDAFKIVRHEGTRTLW---SGLPATLVMTVPATAIYETAYDQLKAFLCGR 155
QY 61 ---SDVTHPGNSHIANGAAGCVATLLHDAAMNPAE 93
Db 156 ALTSDLYAP-----MWAGALARLGTVTVISPLE 183
RESULT 6
DNC_HUMAN
ID DNC_HUMAN STANDARD; PRT; 320 AA.
AC Q9HC21;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial deoxynucleotide carrier (Mitochondrial uncoupling
protein 1).
GN SIC25A19 OR DNC OR MUP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX PubMed=11226231;
RA Dolce V., Piermonte G., Runswick M.J., Palmieri F., Walker J.E.;
RT "The human mitochondrial deoxynucleotide carrier and its role in the
toxicity of nucleoside antivirals.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:2284-2288(2001).
[2]
RP SEQUENCE FROM N.A.
RA Renard S., Mondesert G., Besnard F.;
RT "MUP 1, a mitochondrial uncoupling protein.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Responsible for the uptake of deoxynucleotides into the
matrix of the mitochondria. Transports all four deoxy NDPs, and,
less efficiently, the corresponding dNTPs. Does not transport
dNMPs, NMPs, deoxynucleosides, nucleosides, purines, or
pyrimidines. Supply deoxynucleotides to the mitochondrial matrix
for conversion to triphosphates and incorporation into
mitochondrial DNA.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane.
CC -1- TISSUE SPECIFICITY: Expressed in all tissues examined except for
placenta. Highest levels in colon, kidney, lung, testis, spleen,
and brain.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- DISEASE: Likely to be medically important by providing the means
of uptake into mitochondria of nucleoside analogs, leading to the
mitochondrial impairment that underlies the toxic side effects of
such drugs in the treatment of viral illnesses, including AIDS,
and in cancer therapy.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AJ251857; CAC27560.1; -
DR EMBL; AJ301616; CAC37793.1; -
DR EMBL; AF182404; AAG16903.1; -
DR EMBL; BC001075; AAH01075.1; -
DR EMBL; BC005120; AAH05120.1; -


```
DR MIM; 606521; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
FT SEQUENCE 320 AA; 35511 MW; 57CE0F01D538B1BE CRC64;

Query Match 17.7%; Score 89; DB 1; Length 320;
Best Local Similarity 27.3%; Pred. No. 0.0077;
Matches 24; Conservative 18; Mismatches 28; Indels 18; Gaps 4;

QY 7 DPAARYRNVLALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKKTLSDV 63
Db 54 DPSAKYHGILQASRQLQEGTAFWKGHVPAQLISIGVGAVQFLSF-----EMLTSL 106
QY 64 IHPGG-----NSH-IANGAGCVATL 83
Db 107 VHRGSYIDAREFSVHFVCGGLAACMATL 134

RESULT 7
SFCL_YEAST
ID SFCL_YEAST STANDARD; PRT; 322 AA.
AC P33303;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Succinate/fumarate mitochondrial transporter (Regulator of acetyl-CoA
DE synthetase activity).
OS SFCL OR ACRI OR YJR095W OR J1921.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94203187; PubMed=7908717;
RA Fernandez M., Fernandez E., Rodicio R.;
FT "ACR1, a gene encoding a protein related to mitochondrial carriers,
FT is essential for acetyl-CoA synthetase activity in Saccharomyces
FT cerevisiae.";
RL Mol. Gen. Genet. 242:727-735(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Ramezani Rad M., Kirchath L., Hollenberg C.P.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98055465; PubMed=9395087;
RA Palmieri L., Lasorsa F.M., De Palma A., Palmieri F., Runswick M.J.,
RA Walker J.E.;
FT "Identification of the yeast ACR1 gene product as a succinate-fumarate
FT transporter essential for growth on ethanol or acetate.";
RL FEBS Lett. 417:114-118(1997).
CC -!- FUNCTION: Transports cytoplasmic succinate, derived from
CC isocitrate by the action of isocitrate lyase in the cytosol, into
CC the mitochondrial matrix in exchange for fumarate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial.
CC -!- INDUCTION: BY ETHANOL OR ACETATE AS SOLE CARBON SOURCES. REPRESSED
CC BY GLUCOSE.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z25485; CAA80973.1; -.
DR EMBL; Z49595; CAA89624.1; -.
DR PIR; S36407; S36407.
DR PIR; S43280; S43280.
DR SGD; S0003856; SFC1.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Transmembrane; Transport; Repeat.
FT CONFLICT 300 322 VREHLENLGIEKNDTPKPKLK -> RKGAFQKIWVYSRR
FT MTHQSQSH (IN REF. 1).
SQ SEQUENCE 322 AA; 35340 MW; EC29718A0F5011A7 CRC64;

Query Match 17.2%; Score 86.5; DB 1; Length 322;
Best Local Similarity 29.6%; Pred. No. 0.015; 36; Indels 7; Gaps 2;
Matches 24; Conservative 14; Mismatches 36;

QY 5 QPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLS--- 61
Db 148 EPNAGPKYNNATHAAVTIVKEGVSALYRGVSLTAARQATNOGANFTVYSKLKEFLQNH 207
QY 62 --DVTHPGGNSHIA--NGAAG 78
Db 208 QMDVLPSWETSCIGLISGAIG 228

RESULT 8
MZOM_BOVIN
ID MZOM_BOVIN STANDARD; PRT; 313 AA.
AC P22292;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial 2-oxoglutarate/malate carrier protein (OGCP).
GN SLC25A11 OR SLC20A4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93091249; PubMed=1457818;
RA Iacobazzi V., Palmieri F., Runswick M.J., Walker J.E.;
FT "Sequences of the human and bovine genes for the mitochondrial 2-
FT oxoglutarate carrier.";
RL DNA Seq. 3:79-88(1992).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX TISSUE=Heart.
RX MEDLINE=91105033; PubMed=2271695;
RA Runswick M.J., Walker J.E., Bisaccia F., Iacobazzi V.,
RA Palmieri F.;
FT "Sequence of the bovine 2-oxoglutarate/malate carrier protein:
FT structural relationship to other mitochondrial transport proteins.";
RL Biochemistry 29:11033-11040(1990).
CC -!- FUNCTION: CATALYZES THE TRANSPORT OF 2-OXOGLUTARATE ACROSS THE
CC INNER MITOCHONDRIAL MEMBRANE IN AN ELECTRONEUTRAL EXCHANGE FOR
CC MALATE OR OTHER DICARBOXYLIC ACIDS, AND PLAYS AN IMPORTANT ROLE
CC IN SEVERAL METABOLIC PROCESSES, INCLUDING THE MALATE-ASPARTATE
CC SHUTTLE, THE OXOGLUTARATE/ISOCITRATE SHUTTLE, IN GLUCONEOGENESIS
CC FROM LACTATE, AND IN NITROGEN METABOLISM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- TISSUE SPECIFICITY: HEART, LIVER, AND BRAIN.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
```



```
RT and of a tRNAThr gene." ;
RL Yeast 13:369-372(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (Potential).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99228; CAA67613.1; -.
DR EMBL; 273042; CAA97286.1; -.
DR SGD; S0003489; YGR257C.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 17 36 POTENTIAL.
FT TRANSMEM 124 140 POTENTIAL.
FT TRANSMEM 162 182 POTENTIAL.
FT TRANSMEM 268 286 POTENTIAL.
SQ SEQUENCE 366 AA; 40763 MW; B0358B6EBE818CB1E CRC64;

Query Match 15.5%; Score 78; DB 1; Length 366;
Best Local Similarity 34.0%; Pred. No. 0.16;
Matches 16; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 11 RYRNVEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKIK 57
Db 101 RPNGLTFAPTKIASVEGITSRLWRGSLTLLMAIPANMVYFGYEYIR 147

RESULT 11
ADT_ANOGA
ID ADT_ANOGA STANDARD; PRT; 301 AA.
AC Q27238;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ADP,ATP carrier protein (ADP/ATP translocase) (Adenine nucleotide
DE translocator) (ANT).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=G3;
RX MEDLINE=943348635; PubMed=8069414;
RA Beard C.B., Crews-Oyen A.E., Kumar V.K., Collins F.H.:
RA "A CDNA encoding an ADP/ATP carrier from the mosquito Anopheles
RA gambiae." ;
RL Insect Mol. Biol. 3:35-40(1994).
CC -!- FUNCTION: CATALYZES THE EXCHANGE OF ADP AND ATP ACROSS THE
CC MITOCHONDRIAL INNER MEMBRANE.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11618; AAB04104.1; -.
DR EMBL; L11617; AAB04105.1; -.
DR InterPro; IPR002067; Mit_carrier.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 14 31 1 (POTENTIAL).
FT TRANSMEM 75 93 2 (POTENTIAL).
FT TRANSMEM 119 136 3 (POTENTIAL).
FT TRANSMEM 178 197 4 (POTENTIAL).
FT TRANSMEM 216 233 5 (POTENTIAL).
FT TRANSMEM 275 293 6 (POTENTIAL).
SQ SEQUENCE 301 AA; 32863 MW; 4CC9E17C9F8DA08B CRC64;

Query Match 15.4%; Score 77.5; DB 1; Length 301;
Best Local Similarity 28.2%; Pred. No. 0.15;
Matches 20; Conservative 13; Mismatches 35; Indels 3; Gaps 1;

QY 12 RYRNVEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKIKLSDVIHFEGNSH 71
Db 156 FNGLLDCLKTKVSDGIIGLYRGENVVQGIITIIYAAVFGCFDTAKGLMPLD---PKNTSI 212

QY 72 IANGAGCVAT 82
Db 213 FVSWAIAQVVT 223

RESULT 12
MZOM_MOUSE
ID MZOM_MOUSE STANDARD; PRT; 313 AA.
AC Q9CR62;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial 2-oxoglutarate/malate carrier protein (OGCP).
GN SLC25A11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.W., Staehli F., Suzuki K., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection." ;
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
```

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:19:23 ; Search time 130.92 Seconds
(without alignments)
124.210 Million cell updates/sec

Title: US-09-870-113-10
Perfect score: 502
Sequence: 1 MQSLQDPDPAARYRNVLALW.....GAAGCVATLLHDAAMNPAEG 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_protist.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match %	Description	ID	
1	496	98.8	Q96A46	Q96A46	homo sapien
2	351	69.9	Q9NYZ2	Q9NYZ2	homo sapien
3	348	69.3	Q920G8	Q920G8	mus musculus
4	280	55.8	Q969S1	Q969S1	homo sapien
5	276	55.0	Q912Y0	Q912Y0	mus musculus
6	216.5	43.1	Q23125	Q23125	caenorhabditis
7	171.5	34.2	Q9VAY3	Q9VAY3	rosophila
8	168	33.5	Q94638	Q94638	onchocerca
9	167.5	33.4	Q9NHV6	Q9NHV6	onchocerca
10	166	33.1	Q94634	Q94634	onchocerca
11	135	26.9	Q9LMJ6	Q9LMJ6	arabidopsis
12	130.5	26.0	Q82049	Q82049	ribes nigru
13	130	25.9	Q64731	Q64731	arabidopsis
14	118	23.5	Q14281	Q14281	schizosacch
15	115	22.9	Q921P8	Q921P8	mus musculus
16	114	22.7	Q9CYJ1	Q9CYJ1	mus musculus

17	108	21.5	321	4	Q9BSK2	Q9BSK2	homo sapien
18	101.5	20.2	384	5	Q27244	Q27244	caenorhabditis
19	99.5	19.8	322	10	Q9ARL9	Q9ARL9	hordium vul
20	99.5	19.8	413	5	Q9WVF9	Q9WVF9	rosophila
21	99	19.7	307	4	Q14589	Q14589	homo sapien
22	98	19.5	322	10	Q9AX03	Q9AX03	oryza sativ
23	96	19.1	338	4	Q9UHR1	Q9UHR1	homo sapien
24	95.5	19.0	310	4	Q9NVN5	Q9NVN5	homo sapien
25	95.5	19.0	311	4	Q96CQ1	Q96CQ1	homo sapien
26	95.5	19.0	311	11	Q922G0	Q922G0	mus musculus
27	94.5	18.8	296	10	Q9CA93	Q9CA93	arabidopsis
28	94.5	18.8	313	10	Q9SJY5	Q9SJY5	arabidopsis
29	94.5	18.8	313	10	Q94K32	Q94K32	arabidopsis
30	94	18.7	318	11	Q9DAM5	Q9DAM5	mus musculus
31	93.5	18.6	336	11	Q9D981	Q9D981	mus musculus
32	91	18.1	132	4	Q9H0G8	Q9H0G8	homo sapien
33	90	17.9	363	10	Q9M228	Q9M228	arabidopsis
34	89.5	17.8	337	4	Q9P182	Q9P182	homo sapien
35	89.5	17.8	351	4	Q9Y379	Q9Y379	homo sapien
36	89.5	17.8	351	4	Q9UF66	Q9UF66	homo sapien
37	89.5	17.8	359	4	Q9BZJ4	Q9BZJ4	homo sapien
38	89.5	17.8	359	11	Q9D8K8	Q9D8K8	mus musculus
39	89	17.7	320	4	Q9HC21	Q9HC21	homo sapien
40	88	17.7	348	10	Q9SH98	Q9SH98	arabidopsis
41	88	17.5	358	5	Q9VQ37	Q9VQ37	rosophila
42	88	17.5	365	5	Q9STJ5	Q9STJ5	rosophila
43	87.5	17.4	310	5	Q18844	Q18844	caenorhabditis
44	86.5	17.2	449	5	Q9VI05	Q9VI05	rosophila
45	85	16.9	311	10	Q9C6K8	Q9C6K8	arabidopsis

ALIGNMENTS

RESULT	1
Q96A46	
ID	Q96A46 PRELIMINARY; PRT; 364 AA.
AC	Q96A46:
DT	01-DEC-2001 (TREMREL. 19, Created)
DT	01-DEC-2001 (TREMREL. 19, Last sequence update)
DT	01-DEC-2001 (TREMREL. 19, Last annotation update)
DE	MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE CARRIER SPLICASE VARIANT).
DE	HMRS3/4.
GN	HMRS3/4.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21195335; PubMed=11297739;
RA	Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
RA	Larsson C., Suomalainen A.;
RT	"Characterization of a novel human putative mitochondrial transporter homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.";
RL	FEBS Lett. 494:79-84(2001).
DR	EMBL; AJ303077; CAC27996.1; -
DR	EMBL; AF327402; AAK49519.1; -
SQ	SEQUENCE 364 AA; 39271 MW; 22049C4F8DD7A577 CRC64;

Query Match	98.8%	Score	496;	DB	4;	Length	364;
Best Local Similarity	100.0%	Pred. No.	2e-47;				
Matches	93;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MQSLQDPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60					
Db	100	MQSLQDPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 159					
QY	61	SDVIHPGGNSHTANGAAGCVATLLHDAAMNPAE 93					
Db	160	SDVIHPGGNSHTANGAAGCVATLLHDAAMNPAE 192					

```

RESULT 2
Q9NYZ2 ID Q9NYZ2 PRELIMINARY; PRT; 347 AA.
AC Q9NYZ2: PRELIMINARY; PRT; 347 AA.
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HFO15 PROTEIN.
GN HFO15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human hypothalamus.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223466; AAF64141.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr.3
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN.2.
SQ SEQUENCE 347 AA; 37828 MW; 700DE61B230E001E CRC64;

Query Match 69.9%; Score 351; DB 4; Length 347;
Best Local Similarity 67.7%; Pred. No. 3.3e-31;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLQPPDPAARYNRVLEALWRIIRTEGLWRPMTGLNTATGAGPAHALYFACYEKLKRTL 60
Db 73 MOSLSPDPKQYTSIYGALKKIMRTGFWRLRGVNVIMGAGPAHAMFYACIENMKRTL 132

QY 61 SDVIHPGNSHIANGAGCVATLLHDAAMPAP 93
Db 133 NDVPHQGNHSLANGVAGSMATLLHDAVMNPAP 165

RESULT 3
Q920G8 ID Q920G8 PRELIMINARY; PRT; 338 AA.
AC Q920G8: PRELIMINARY; PRT; 338 AA.
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RA Li Q.-Z., Ruan Q.-G., Eckenrode S., Shi J.-D., Cruz P., Wang C.-Y.,
RT She J.-X.;
RL "A new gene which is highly expressed in NOD mice spleen.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF288621; AAL23859.1; -.
SQ SEQUENCE 338 AA; 37510 MW; BB35B1F70C56A3FE CRC64;

Query Match 69.3%; Score 348; DB 11; Length 338;
Best Local Similarity 68.8%; Pred. No. 7e-31;
Matches 64; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 1 MOSLQPPDPAARYNRVLEALWRIIRTEGLWRPMTGLNTATGAGPAHALYFACYEKLKRTL 60
Db 73 MOSLSPDPKQYTSIYGALKKIMRTGFWRLRGVNVIMGAGPAHAMFYACIENMKRTL 132

QY 61 SDVIHPGNSHIANGAGCVATLLHDAAMPAP 93
Db 133 NDVPHQGNHSLANGVAGSMATLLHDAVMNPAP 165

```

```

RESULT 4
Q969S1 ID Q969S1 PRELIMINARY; PRT; 155 AA.
AC Q969S1: PRELIMINARY; PRT; 155 AA.
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MITOCHONDRIA SOLUTE CARRIER PROTEIN (HYPOTHETICAL 16.8 KDA
DE PROTEIN).
GN MSCP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Li Q., Eckenrode S., Ruan Q., Wang C., Shi J., McIndoe R.A., She J.;
RT "Molecular cloning of a novel mitochondrial solute carrier protein
RT (MSCP) gene from mouse and human and its down-regulation in mouse
RT spleen during the maturation of the immune system.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RA TISSUE=UTERUS, AND LEIOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY032628; AAK38154.1; -.
DR EMBL; BC015013; AAH15013.1; -.
KW Hypothetical protein.
SQ SEQUENCE 155 AA; 16832 MW; 8ACB98A483C8E6EF CRC64;

Query Match 55.8%; Score 280; DB 4; Length 155;
Best Local Similarity 65.3%; Pred. No. 1.2e-23;
Matches 49; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 MOSLQPPDPAARYNRVLEALWRIIRTEGLWRPMTGLNTATGAGPAHALYFACYEKLKRTL 60
Db 73 MOSLSPDPKQYTSIYGALKKIMRTGFWRLRGVNVIMGAGPAHAMFYACIENMKRTL 132

QY 61 SDVIHPGNSHIANG 75
Db 133 NDVPHQGNHSLANG 147

RESULT 5
Q91ZY0 ID Q91ZY0 PRELIMINARY; PRT; 182 AA.
AC Q91ZY0: PRELIMINARY; PRT; 182 AA.
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MITOCHONDRIAL CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RA Li Q., Eckenrode S., Wang C., Ruan Q., Shi J., McIndoe R.A., She J.;
RT "A novel mouse mitochondrial carrier protein gene is up-regulated from
RT young to adult NOD mice.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF361699; AAL27990.1; -.
SQ SEQUENCE 182 AA; 19838 MW; 18E2C5E801228693 CRC64;

Query Match 55.0%; Score 276; DB 11; Length 182;
Best Local Similarity 66.7%; Pred. No. 4e-23;
Matches 50; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

```



```

QY 1 MOSLPDPAARYNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
    ||||| ||| ||| :: ||||| ||| ||| ||| ||||| ||||| ||||| |||||
DB 73 MOSLNPDPKARTSYIGALKRIMHTEGFWRLRGLNVMGAGPAHAMFYFACIENKRTL 132
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 SDVIHPGGNSHIANG 75
    :|| ||||| |||
DB 133 NDVESHQGNHSLANG 147
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
Q23125 PRELIMINARY; PRT; 312 AA.
AC Q23125;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE W02B12.9 PROTEIN.
GN W02B12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN RA Swinburne J., Ainscough R.;
RP SEQUENCE FROM N.A.
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RT none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z66521; CA91399.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN2.
SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;

Query Match 43.1%; Score 216.5; DB 5; Length 312;
Best Local Similarity 50.0%; Pred. No. 3.5e-16;
Matches 47; Conservative 16; Mismatches 28; Indels 3; Gaps 2;

QY 1 MOSLPDPAARYNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
    ||||| ||| ||| :: ||||| ||| ||| ||| ||||| ||||| ||||| |||||
DB 45 MOSLCPCPETKCPPTVHSLMSIVKREGWRLRPLRGYNVAAGSMPAHALYFTVYERKGYL 104
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 SDVIHPGGNSH--IANGAAGCVATLLHDAAMNPAE 93
    : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
DB 105 TG--NSAGHSNTLAYGASGVVATLLHDAIMNPAE 136
    : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

RESULT 7
Q9VAY3 PRELIMINARY; PRT; 379 AA.
AC Q9VAY3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CG4963 PROTEIN (CH09840P).
GN CG4963.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN RA Swinburne J., Ainscough R.;
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=I, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC -!- INNER MEMBRANE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; AE003763; AAF56764.1; -.
DR EMBL; AY060268; AAL25307.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carr; 3.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 34.2%; Score 171.5; DB 5; Length 379;
Best Local Similarity 40.9%; Pred. No. 4.8e-11;
Matches 38; Conservative 16; Mismatches 34; Indels 5; Gaps 2;

QY 1 MOSLPDPAARYNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
    ||||| ||| ||| :: ||||| ||| ||| ||| ||||| ||||| ||||| |||||
DB 42 MQSL--SPPTKNMNVSTLRMTREGLLRPIRGASAVVLGAGPAHSLYFAAYETKLT 99
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 SDVIHPGGNSHIANGAAGCVATLLHDAAMNPAE 93
    : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
DB 100 AKFTSVRNINLVISGA---VATLIHDAISSPTD 129
    : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

```

```

RESULT 8
Q94638 ID Q94638 PRELIMINARY; PRT; 303 AA.
AC Q94638;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL mitochondrial solute carriers."
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45998; AAB19037.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 303 AA; 33861 MW; 224BFA347E5D617C CRC64;

Query Match 33.5%; Score 168; DB 5; Length 303;
Best Local Similarity 36.6%; Pred. No. 9.2e-11;
Matches 34; Conservative 20; Mismatches 37; Indels 2; Gaps 1;

QY 1 MQSLQDPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 42 LQSLPCPCPTSCPTAMHSLMSVMKREGLRLSLKGVNAVLTIPAHAFYTYVYENS KAYL 101
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 SDVIHPCGNSHIANGAAGCVATLLHDAAMPAPAE 93
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 102 LN--NPRVSNMSYATISGALATVIHDAVMNPAPAE 132
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
Q9NHV6 ID Q9NHV6 PRELIMINARY; PRT; 380 AA.
AC Q9NHV6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOPHETICAL 41.8 KDA PROTEIN.
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Schmid K.J., Tautz D.;
RT "A screen for rapidly evolving genes from Drosophila.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217402; AAF73387.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 41844 MW; 4C9AA524B97F8C6C CRC64;

Query Match 33.4%; Score 167.5; DB 5; Length 380;
Best Local Similarity 39.8%; Pred. No. 1.4e-10;

```

```

Matches 37; Conservative 16; Mismatches 35; Indels 5; Gaps 2;

QY 1 MQSLQDPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 42 MQSL--SPPTKNNIVSTLRTMITREGLLRPIRGASAVVLGAGPTSLYFAAVEMTKELT 99
: : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 SDVIHPCGNSHIANGAAGCVATLLHDAAMPAPAE 93
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 100 AKFTSVRNLNVTISGA---VATLIHDAISSPTD 129
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q94634 ID Q94634 PRELIMINARY; PRT; 301 AA.
AC Q94634;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca gibsoni.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6284;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL mitochondrial solute carriers."
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45997; AAB19036.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 301 AA; 34176 MW; C0BA8D819FB8EA79 CRC64;

Query Match 33.1%; Score 166; DB 5; Length 301;
Best Local Similarity 36.6%; Pred. No. 1.5e-10;
Matches 34; Conservative 20; Mismatches 37; Indels 2; Gaps 1;

QY 1 MQSLQDPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 42 LQSLPCPCPTSCPTAMHSLMSVMKREGLRLSLKGVNAVLTIPAHAFYTYVYENS KAYL 101
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 SDVIHPCGNSHIANGAAGCVATLLHDAAMPAPAE 93
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 102 LN--NPRVSNMSYATISGALATVIHDAVMNPAPAE 132
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Q9LMJ6 ID Q9LMJ6 PRELIMINARY; PRT; 781 AA.
AC Q9LMJ6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE F10K1.26 PROTEIN.
GN F10K1.26.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu S.X., Chan A., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
RA Sakano H., Toriumi M., Chung M., Goldsmith A., Liu A., Liu A.,
RA Vaysberg M., Altati H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "The sequence of BAC F10K1 from Arabidopsis thaliana chromosome 1.";

```



```
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; Z99168; CAB16300.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
SQ SEQUENCE 303 AA; 32652 MW; 69975CDE18107AB5 CRC64;

Query Match 23.5%; Score 118; DB 3; Length 303;
Best Local Similarity 33.3%; Pred. No. 3.7e-05;
Matches 26; Conservative 17; Mismatches 31; Indels 4; Gaps 1;

QY 14 NVLEALWRIIRTEGLWRPVRGRLNVTATGAGPAHALYFACYEKLKKTLSDVHPGNSHIA 73
DB 60 NTVNSVIKISSSTEGYSLWRGSISSVIMGAGPSHAIFYSVLEFFKSK- ---INASPDRLA 115
QY 74 NGAAGCVATLLHDAAMNP 91
DB 116 SALAGACAITSDAFWTP 133

RESULT 15
Q921P8
ID Q921P8 PRELIMINARY; PRT; 320 AA.
AC Q921P8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO RIKEN CDNA 5730438N18 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011293; AAH11293.1; -.
SQ SEQUENCE 320 AA; 35050 MW; F40B5912706583CD CRC64;

Query Match 22.9%; Score 115; DB 11; Length 320;
Best Local Similarity 32.9%; Pred. No. 8.5e-05;
Matches 28; Conservative 12; Mismatches 41; Indels 4; Gaps 2;

QY 8 PAARYNVLEALWRIIRTEGLWRPVRGRLNVTATGAGPAHALYFACYEKLKKTLSDVHPG 67
DB 67 PTVPTPGLLQVLKSILEKEGPKSLFRGLGNLVGVAPSRVAFACYSRAKEQNGVFVPN 126
QY 68 GNS-HIANGACGCVATLLHDAAMNP 91
DB 127 SNTVHLSAGSAAFVT---NTLMNP 148
```

Search completed: August 27, 2002, 04:19:24
Job time: 1291 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:56:00 ; Search time 138.54 Seconds

(without alignments)
105.029 Million cell updates/sec

Title: US-09-870-113-12

Perfect score: 695

Sequence: 1 MQSLQPDPAARYNRVLEALW.....QFAESTSVLGVNSVTIFYH 131

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	505	72.7	289	21 AAB50383	Human uncoupling p
2	505	72.7	289	22 AAB60113	Human transport pr
3	505	72.7	318	22 AAM41505	Human polypeptide
4	476	68.5	268	22 AAM39719	Human polypeptide
5	352	50.6	299	22 AAM79039	Human protein SPQ
6	351	50.5	272	21 AAB42980	Human ORFX ORF2744
7	351	50.5	331	22 AAB60658	Human mitochondria
8	351	50.5	366	22 AAM00938	Human bone marrow
9	322	46.3	677	22 AAU29748	Novel human secret
10	276	39.7	155	21 AAB42966	Human ORFX ORF2730
11	178.5	25.7	379	22 ABB61130	Drosophila melanog

12	165.5	23.8	59	22	AAM85891	Human immune/haema
13	141.5	20.4	289	18	AAW17054	Blackcurrant RIB7
14	139	20.0	278	21	AAG22079	Arabidopsis thalia
15	139	20.0	278	21	AAG43094	Arabidopsis thalia
16	139	20.0	290	21	AAG22078	Arabidopsis thalia
17	139	20.0	290	21	AAG43093	Arabidopsis thalia
18	139	20.0	331	21	AAG22077	Arabidopsis thalia
19	139	20.0	331	21	AAG43092	Arabidopsis thalia
20	121	17.4	130	22	AAM00825	Human bone marrow
21	99.5	14.3	413	22	ABP69108	Drosophila melanog
22	97	14.0	241	21	AAG07172	Arabidopsis thalia
23	97	14.0	289	21	AAG07171	Arabidopsis thalia
24	97	14.0	296	21	AAG07170	Arabidopsis thalia
25	95.5	13.7	310	22	AAG93001	Human protein sequ
26	95.5	13.7	311	21	AAB50382	Human uncoupling p
27	95.5	13.7	311	22	AAM39173	Human polypeptide
28	95.5	13.7	374	22	AAM40959	Human polypeptide
29	95.5	13.7	374	22	AAM40960	Human polypeptide
30	95	13.7	308	21	AAB25780	Human secreted pro
31	95	13.7	308	22	AAB80951	Human carnitine ca
32	95	13.7	308	22	AAB75366	Human secreted pro
33	94.5	13.6	237	21	AAG19645	Arabidopsis thalia
34	94.5	13.6	237	21	AAG50519	Arabidopsis thalia
35	94.5	13.6	284	21	AAG19644	Arabidopsis thalia
36	94.5	13.6	284	21	AAG50518	Arabidopsis thalia
37	94.5	13.6	313	21	AAG19643	Arabidopsis thalia
38	94.5	13.6	313	21	AAG50517	Arabidopsis thalia
39	94.5	13.6	351	21	AAB50378	Human uncoupling p
40	94.5	13.6	365	21	AAB42783	Human ORFX ORF2547
41	92	13.2	96	21	AAG27453	Arabidopsis thalia
42	92	13.2	121	21	AAG27452	Arabidopsis thalia
43	90	12.9	243	21	AAG27858	Arabidopsis thalia
44	90	12.9	285	21	AAG27857	Arabidopsis thalia
45	90	12.9	308	22	AAM38891	Human polypeptide

ALIGNMENTS

RESULT 1

AAB50383
ID AAB50383 standard; Protein; 289 AA.

XX
AC AAB50383;

XX
DT 12-MAR-2001 (first entry)

XX
DE Human uncoupling protein #6.

XX
DE Human; uncoupling protein; immunosuppressive; antiarthritic;

KW antirheumatic; antiproliferative; cardiac; vasotropic;

KW cerebroprotective; neuroprotective, antibacterial; ophthalmological;

KW gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;

KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;

KW infertility.

XX
OS Homo sapiens.

XX
PN WO200061614-A2.

XX
PD 19-OCT-2000.

XX
PF 06-APR-2000; 2000WO-US09534.

XX
PR 09-APR-1999; 99US-0128701.

PR 08-JUL-1999; 99US-0142821.

PR 18-AUG-1999; 99US-0149448.

PR 12-NOV-1999; 99US-0164751.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;

DR WPI: 2000-656322/63.
XX N-PSDB; AAC90457.
XX
PT Uncoupling proteins and nucleic acid sequences encoding them, useful
PT for detecting, preventing and treating proliferative, neurological,
PT immune system, cardiovascular and gastrointestinal disorders -
XX
PS Claim 11; Page 323-324; 343pp; English.
XX
XX The present sequence is a human uncoupling protein. The nucleotide
CC sequences encoding the uncoupling proteins may be used for
CC the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
CC or research purposes. The uncoupling protein encoded by the nucleotide
CC sequences may be used to treat disorders such as neural, immune,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders, wounds, infectious diseases,
CC thrombosis, arthritis, and infertility.
XX
SQ Sequence 289 AA;

Query Match 72.7%; Score 505; DB 21; Length 289;
Best Local Similarity 87.5%; Pred. No. 5.5e-51;
Matches 98; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

QY 1 MQSLQDPDAARYNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKIKKTL 60
Db 25 mqslqpdaarynvlealwriirteglwrp mrglnvtatgagpahalyfacyekikktl 84

QY 61 SDVIHFGGNSHIANGAAGCVATLLHDAAMNPAE-----GNDSSTYHSVGSC 106
Db 85 sdvihpggnshiangaagcvatllhdaamnpaevvkgrmqmynspyrhrvtcd 136

RESULT 2
AAB60113
ID AAB60113 standard; Protein; 289 AA.
XX
XX AAB60113;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33.
XX
XX Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer.
XX
XX Homo sapiens.
XX
XX WO200078953-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US16668.
XX
XX 17-JUN-1999; 99US-0139923.
XX 10-AUG-1999; 99US-0148177.
XX 18-AUG-1999; 99US-0149357.
XX 28-OCT-1999; 99US-0162287.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI: 2001-041424/05.
DR N-PSDB; AAF27733.
XX
XX Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -

XX Claim 2; Page 133-134; 165pp; English.
XX
XX The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
SQ Sequence 289 AA;

Query Match 72.7%; Score 505; DB 22; Length 289;
Best Local Similarity 87.5%; Pred. No. 5.5e-51;
Matches 98; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

QY 1 MQSLQDPDAARYNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKIKKTL 60
Db 25 mqslqpdaarynvlealwriirteglwrp mrglnvtatgagpahalyfacyekikktl 84

QY 61 SDVIHFGGNSHIANGAAGCVATLLHDAAMNPAE-----GNDSSTYHSVGSC 106
Db 85 sdvihpggnshiangaagcvatllhdaamnpaevvkgrmqmynspyrhrvtcd 136

RESULT 3
AAM41505
ID AAM41505 standard; Protein; 318 AA.
XX
XX AAM41505;
XX
DT 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 6436.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW anyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR N-PSDB; AA160661.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6436; 10078pp; English.
XX

(HYSE-) HYSEQ INC.
Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
Xue AJ, Yang Y, Wehrman T, Goodrich R;
WPI; 2001-476283/51.
N-PSDB; AAK52172.
Nucleic acids encoding polypeptides with cytokine-like activities,
useful in diagnosis and gene therapy -
Claim 20; Page 4043; 6221pp; English.
The invention relates to polynucleotides (AAK51456-AAK53435) and the
encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation.
Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
(AAM80020) are omitted as the relevant pages from the sequence listing
were missing at the time of publication.
Sequence 299 AA;
Query Match 50.6%; Score 352; DB 22; Length 299;
Best Local Similarity 68.8%; Pred. No. 6.1e-33;
Matches 64; Conservative 12; Mismatches 17; Indels 0; Gaps 0;
QY 1 MQSLQDPAAARNVLEALWRIIRTEGLWRPMRGVNVLTATGAGPAHALYFACYEKLKRTL 60
Db 100 mqlspdpkaqytsygalcklmrtgfwprlgrgvnmimgagpahamyfacyenmkrtl 159
QY 61 SDVHPGNSHIANGAAGCVATLLHDAAMNPAE 93
Db 160 ndvfhhqgnshlangiagsmatlhlhdavmmpae 192
RESULT 6
AAB42980
ID AAB42980 standard; Protein: 272 AA.
AC AAB42980;
DT 08-FEB-2001 (first entry)
DE Human ORFX ORF2744 polypeptide sequence SEQ ID NO:5488.
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipariatic; antiparkinsonian; neotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX Homo sapiens.
OS WO200058473-A2.
PN

XX PD 05-OCT-2000.
XX PF 31-MAR-2000; 2000WO-US08621.
XX PR 31-MAR-1999; 99US-0127607.
XX PR 02-APR-1999; 99US-0127636.
XX PR 05-APR-1999; 99US-0127728.
XX PR 30-MAR-2000; 2000US-0540763.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX DR WPI; 2000-602362/57.
XX DR N-PSDB; AAC77189.
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX Claim 11; Page 4662-4663; 5507pp; English.
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipariatic; antiparkinsonian; neotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX Sequence 272 AA;
SQ
Query Match 50.5%; Score 351; DB 21; Length 272;
Best Local Similarity 67.7%; Pred. No. 7.1e-33;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;
QY 1 MQSLQDPAAARNVLEALWRIIRTEGLWRPMRGVNVLTATGAGPAHALYFACYEKLKRTL 60
Db 7 mqlspdpkaqytsygalcklmrtgfwprlgrgvnmimgagpahamyfacyenmkrtl 66
QY 61 SDVHPGNSHIANGAAGCVATLLHDAAMNPAE 93
Db 67 ndvfhhqgnshlangiagsmatlhlhdavmmpae 99
RESULT 7
AAB60658
ID AAB60658 standard; Protein: 331 AA.
XX AAB60658;
AC AAB60658;
XX 04-MAY-2001 (first entry)
DT Human mitochondrial solute carrier protein hmSC-o.
DE Human mitochondrial solute carrier protein hmSC-o.
XX Human mitochondrial solute carrier protein; hmSC-o; hypothalamus;
KW preparation; detection.


```

XX OS Homo sapiens.
XX CN1269409-A.
XX PD 11-OCT-2000.
XX PF 17-MAR-2000; 2000CN-0114958.
XX PR 17-MAR-2000; 2000CN-0114958.
XX PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
XX PI Zhang X, Gao X, Xiao H;
XX WPI; 2001-050544/07.
XX DR N-PSDB; AAF59920.
XX PT New human mitochondrion solute carrier protein and its nucleic acid -
XX PS Claim 4; Page 20; 21pp; Chinese.
XX CC The invention relates to a novel human mitochondrial solute carrier
XX CC protein, hMSC-o (AAB6058), and cDNA encoding it (AAF59920). hMSC-o is
XX CC expressed in normal human hypothalamus tissue. The invention also relates
XX CC to the preparation of hMSC-o proteins and nucleic acids, and the
XX CC detection of hMSC-o proteins and nucleic acids in a sample. The present
XX CC sequence represents hMSC-o.
XX SQ Sequence 331 AA;

Query Match 50.5%; Score 351; DB 22; Length 331;
Best Local Similarity 67.7%; Pred. No. 9.2e-33;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLOPDAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
Db mqlslpdkagqtslygalkkkmrtgfwprlgvnmimgagpahamfacyemmkrtl 116
QY 61 SDVIHPGNSHIANGACGCVATLLHDAAMNPAE 93
Db 117 ndvfhhqgnshlangiagdsmatllhdavmnpae 149

RESULT 8
AAM00938
ID AAM00938 standard; Protein; 366 AA.
XX AC AAM00938;
XX DT 01-OCT-2001 (first entry)
XX DE Human bone marrow protein, SEQ ID NO: 414.
XX KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
XX KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
XX KW immunosuppressive; gene therapy; cytokine cell proliferation;
XX KW cell differentiation modulator; immune disorder; infection; cancer;
XX KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX OS Homo sapiens.
XX WO200153453-A2.
XX PN 26-JUL-2001.
XX PD 23-DEC-2000; 2000WO-US34960.
XX PF 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.

```

```

PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Drmanac RT;
XX WPI; 2001-488707/53.
XX DR N-PSDB; AAF90057.
XX CC Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX CC for treating e.g. cancer and immune deficiency disorders -
XX PS Claim 10; Page 504-505; 648pp; English.
XX CC The present sequence is one of 251 novel human polypeptides encoded
XX CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
XX CC polypeptide encoded by it are useful in the treatment of various
XX CC immune deficiencies and disorders. The deficiencies and disorders may
XX CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
XX CC infection, or may result from an autoimmune disorder, a coagulation
XX CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX CC suppression of an inflammatory response or treatment of a nervous
XX CC system disorder such as Alzheimer's disease. Detection of the presence
XX CC or increased expression of the polynucleotide or the protein it
XX CC encodes is useful for the diagnosis and/or prognosis of one
XX CC or more types of cancer. The polynucleotide and polypeptide can be
XX CC used as nutritional sources or supplements and in the screening of
XX CC chemical compounds as potential drugs.
XX SQ Sequence 366 AA;

Query Match 50.5%; Score 351; DB 22; Length 366;
Best Local Similarity 67.7%; Pred. No. 1.1e-32;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLOPDAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
Db mqlslpdkagqtslygalkkkmrtgfwprlgvnmimgagpahamfacyemmkrtl 160
QY 61 SDVIHPGNSHIANGACGCVATLLHDAAMNPAE 93
Db 161 ndvfhhqgnshlangiagdsmatllhdavmnpae 193

RESULT 9
AAU29748
ID AAU29748 standard; Protein; 677 AA.
XX AC AAU29748;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #239.
XX KW Human; vaccination; gene therapy; nutritional supplement;
XX KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX WO200179449-A2.
XX PN 25-OCT-2001.
XX PD 16-APR-2001; 2001WO-US08656.
XX PF 18-APR-2000; 2000US-0552929.

```


KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
XX WO200171042-A2.
XX
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI N-PSDB; ABL05233.
DR
DR
XX
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
XX Disclosure; SEQ ID NO 10182; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 379 AA;
SQ

Query Match 25.7%; Score 178.5; DB 22; Length 379;
Best Local Similarity 39.3%; Pred. No. 2.4e-12;
Matches 44; Conservative 17; Mismatches 40; Indels 11; Gaps 3;

QY 1 MOSLQDPDPAARYNNVLEALWRIIRTEGLWPMRGLNVTATGAGPAHALYFACYEKLKTL 60
||||| : : : : : ||| : : : : : ||||| : : : : :
Db 42 mgs1--sptkmmnivistlrmittregllrpigasavvlgaqpahslyfaayemtkelt 99
QY 61 SDVIHPGGNSHIANGAGCVATLLHDAANPAE-----GNDSTYHVSVC 106
: : : : : ||||| : : : : : ||| : : : : :
Db 100 akftsrninyvisga---vatlihdaispdtvikgrmqmynspytvsvc 148

RESULT 12
AAM85891
ID AAM85891 standard; Protein; 59 AA.
XX
XX
XX AAM85891;
XX
XX
DT 07-NOV-2001 (first entry)
DE
DE Human immune/haematopoietic antigen SEQ ID NO:13484.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX
XX 09-AUG-2001.
PD

XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 27-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.
XX 14-SEP-2000; 2000US-0232399.
XX 14-SEP-2000; 2000US-0232400.
XX 14-SEP-2000; 2000US-0232401.
XX 14-SEP-2000; 2000US-0233063.
XX 14-SEP-2000; 2000US-0233064.
XX 14-SEP-2000; 2000US-0233065.
XX 21-SEP-2000; 2000US-0234223.
XX 21-SEP-2000; 2000US-0234274.
XX 25-SEP-2000; 2000US-0234997.
XX 25-SEP-2000; 2000US-0234998.
XX 26-SEP-2000; 2000US-0235484.
XX 27-SEP-2000; 2000US-0235834.
XX 27-SEP-2000; 2000US-0235836.
XX 29-SEP-2000; 2000US-0236327.
XX 29-SEP-2000; 2000US-0236367.
XX 29-SEP-2000; 2000US-0236368.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 22-MAY-1999; 99US-0135625.
PR 25-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:58:53 ; Search time 48.11 Seconds
(without alignments)
66.509 Million cell updates/sec

Title: US-09-870-113-12

Perfect score: 695

Sequence: 1 MQSLQDPDPAARYNVLALW.....QFAEESTSVLGVNSVTILFYH 131

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141.5	20.4	289	US-09-068-140A-10	Sequence 10, Appl
2	141.5	20.4	328	US-09-068-140A-15	Sequence 15, Appl
3	94.5	13.6	351	US-08-933-750C-19	Sequence 19, Appl
4	94.5	13.6	351	US-09-234-613-19	Sequence 19, Appl
5	90	12.9	320	US-08-933-750C-12	Sequence 12, Appl
6	90	12.9	320	US-09-234-613-12	Sequence 12, Appl
7	79	11.4	293	US-08-628-291-4	Sequence 4, Appl
8	79	11.4	293	US-09-128-722-4	Sequence 4, Appl
9	79	11.4	317	US-08-628-291-12	Sequence 12, Appl
10	79	11.4	317	US-09-128-722-12	Sequence 12, Appl
11	71	10.2	1094	US-09-268-347-32	Sequence 32, Appl
12	71	10.2	2584	US-08-936-135-4	Sequence 4, Appl
13	69	9.9	306	PCT-US94-09799-1	Sequence 1, Appl
14	68.5	9.9	311	US-08-775-009-32	Sequence 32, Appl
15	68.5	9.9	312	US-09-142-565-2	Sequence 2, Appl
16	68.5	9.9	658	US-08-409-995-5	Sequence 5, Appl
17	68.5	9.9	658	US-08-685-467-5	Sequence 5, Appl
18	68.5	9.9	658	US-08-913-942-5	Sequence 5, Appl
19	68.5	9.9	1098	US-08-409-995-2	Sequence 2, Appl
20	68.5	9.9	1098	US-08-685-467-2	Sequence 2, Appl
21	68.5	9.9	1098	US-09-377-155-32	Sequence 32, Appl
22	68.5	9.9	1098	US-08-913-942-2	Sequence 2, Appl
23	68.5	9.9	1098	US-09-669-974-32	Sequence 32, Appl
24	68.5	9.9	1098	US-09-268-347-44	Sequence 44, Appl
25	67	9.6	432	US-08-937-466-4	Sequence 4, Appl
26	67	9.6	432	US-09-172-528-4	Sequence 4, Appl
27	67	9.6	432	US-09-318-199-4	Sequence 4, Appl

28 67 9.6 432 4 US-09-503-579-4 Sequence 4, Appl
29 66 9.5 2594 4 US-08-718-388-7 Sequence 7, Appl
30 66 9.5 5405 4 US-08-718-388-9 Sequence 9, Appl
31 65.5 9.4 311 2 US-08-775-009-33 Sequence 33, Appl
32 65 9.4 447 4 US-09-160-119-2 Sequence 2, Appl
33 65 9.4 674 4 US-09-160-119-2 Sequence 2, Appl
34 63.5 9.1 298 3 US-08-961-871-10 Sequence 10, Appl
35 63 9.1 308 2 US-08-937-466-2 Sequence 2, Appl
36 63 9.1 308 2 US-09-172-528-2 Sequence 2, Appl
37 63 9.1 308 3 US-09-318-199-2 Sequence 2, Appl
38 63 9.1 308 4 US-09-503-579-2 Sequence 2, Appl
39 63 9.1 571 2 US-08-892-770-5 Sequence 5, Appl
40 62 8.9 469 4 US-09-188-930-339 Sequence 339, Appl
41 62 8.9 888 4 US-09-268-140-4 Sequence 4, Appl
42 61.5 8.8 119 1 US-08-053-171-14 Sequence 14, Appl
43 61.5 8.8 119 1 US-08-053-171-17 Sequence 17, Appl
44 61.5 8.8 119 3 US-08-815-190A-13 Sequence 13, Appl
45 61.5 8.8 2588 3 US-08-936-135-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-068-140A-10
; Sequence 10, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-10

RESULT 5
US-08-933-750C-12
; Sequence 12, Application US/08933750C
; Patent No. 5932442

APPLICANT: Hillman, Jennifer L.
 APPLICANT: Bandman, Olga
 APPLICANT: Shah, Purvi
 APPLICANT: Au-Young, Janice
 APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 APPLICANT: Corley, Neil C.
 TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
 NUMBER OF SEQUENCES: 98
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/933.750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SPLNOT02
CLONE: 207452
US-08-933-750C-12

Query Match 12.9%; Score 90; DB 2; Length 320;
Best Local Similarity 24.6%; Pred. No. 0.0074;
Matches 30; Conservative 25; Mismatches 33; Indels 34;

Db	54	DPSAKYHGIIQASRQILQEEGPTAFWKGHVPQAIIISIGYGAQVQFLSF	EMLTPE
Qy	64	IPHGG-----NSH-IANGAAGCVATLLHDAAMP-----AEGND-----SSTYHS	
Db	107	VHRGSVVDAREFSVHFVCGGLAACMATL-----TVHPVDVLRTFRFAAQGEKPVNTLREA	
Qy	104	GS 105	
Db	163	GT 164	

```

Query Match          12.9%; Score 90; DB 4; Length 320;
Best Local similarity 24.6%; Pred. NO. 0.0074;
Matches 30; Conservative 25; Mismatches 33; Indels 34; Gaps

Qy      7  DPAARYRNVLLEALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACVEKLLKLTSDV 63
         ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||
Db      54  DPSAKYHGIIQAASRQILQEEGPTAFWKGHVHPAQIILSIGYGAQVQLSF-----EMLTET 106

Qy      64  IHPGG-----NSH-IANGAACCVATLLHDAAMNP-----REGND---SSYTHSV 103
         ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||
Db      107  VHRGSVVDAREFSVHFVCGGLAACMATL-----TVHPVDVLTFRFAAQGEEKVYNTLRHAV 162

Qy      104  GS 105
         ||:
Db      163  GT 164

RESULT 7

```

RESULT 7

```

: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
:
: NAME: TMIAN, SUSAN J.
: REGISTRATION NUMBER: 34,103
: REFERENCE/DOCKET NUMBER: 20894/150
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: 716-263-1636
: TELEFAX: 716-263-1600
:
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:

```

QY 51 ACYEKLKKTLSVDVIHPGGNSHIANGAAGCVATLLHDAAMNPAEGNDSSSTYHSVGSCTCIS 110

Db 80 AQVETLFRSINGVLLPGGANLTHSGYSRVAKIFFTKAL---ESFDNGDFFPVWG-TCLG 135
QY 111 LQFAEESTSVLVGN 124
Db 136 L-----EELSVLVSN 145

RESULT 11
US-09-268-347-32
; Sequence 32, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 1094
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-32

Query Match 10.2%; Score 71; DB 4; Length 1094;
Best Local Similarity 27.9%; Pred. No. 8.7;
Matches 29; Conservative 17; Mismatches 46; Indels 12; Gaps 4;
QY 34 GLNVATGAGPAHALYFACYEKL---KKTLSVDVHPGGNSHIANGAAGCVATLLHDAAM 90
Db 150 GVQVTSSENGKAIYFALAKLDMRTATVSDLTITGGSTTTGSATPKYVNTSTASGLN 209
QY 91 PAEG----NDSSYHVSVCSTCISLQFAESTSV---LVGNSTV 127
Db 210 FARGATGANGDTVHLTNIAS--TLQDTLLNTGVWSKLDNGKIT 251

RESULT 12
US-08-936-135-4
; Sequence 4, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-936-135-4

Query Match 10.2%; Score 71; DB 3; Length 2584;
Best Local Similarity 29.1%; Pred. No. 29;
Matches 32; Conservative 15; Mismatches 35; Indels 28; Gaps 6;
QY 3 SLOPDP-AARYRVNLEALWRIIR-TEGLWRPMBGLNVNTATGAGPAHALYFACYEKLAKTL 60
Db 1320 ALERPRSRERPHISPRTRYRILEASNGLTRPLEGLNVAL--ASPLEGLY-ASPGLLYSIL 1376
QY 61 SDVIHPGG-----NSHIANGAAGCVATLLHDAAM 89
Db 1377 EVALARGGLYVALILEITLEGNGLYGLYLSHIS--ARGGLASNLYSVAL 1424

RESULT 13
PCT-US94-09799-1
; Sequence 1, Application PC/TUS9409799
; GENERAL INFORMATION:
; APPLICANT: Kagan, David
; TITLE OF INVENTION: Method and Composition for Weight Reduction
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Watov & Kipnes, P.C.
; STREET: 186 Princeton-Hightstown Rd, PO Box 247
; CITY: Princeton Junction
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08550
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: DOS 6.0
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09799
; FILING DATE: 29-AUG-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,179
; FILING DATE: 25 May 1994
; APPLICATION NUMBER: 08/114,313
; FILING DATE: 30 August 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kipnes, Allen R.
; REGISTRATION NUMBER: 28,433
; REFERENCE/DOCKET NUMBER: 489.1.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-243-0330
; TELEFAX: 609-275-1010
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; TOPOLOGY: Linear
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL TYPE: Brown adipose
; FEATURE:
; OTHER INFORMATION: Cassard, A.M. et al. publication
; OTHER INFORMATION: Genbank - Locus 37607
; PUBLICATION INFORMATION:
; AUTHORS: Cassard, A.M. et al.
; TITLE: Human Uncoupling Protein Gene: Structure,
PCT-US94-09799-1

Result No.	Query	Score	Match	Length	DB	ID	Description
1	216.5	31.2	312	2	T26089	hypothetical prote	
2	143.5	20.6	781	2	A86205	hypothetical prote	
3	139	20.0	331	2	T00582	probable mitochond	
4	125.5	18.1	314	2	S55179	mRNA splice defec	
5	119	17.1	303	2	T39149	probable RNA splic	
6	106	15.3	304	2	S13533	mRNA splice defec	
7	101.5	14.6	384	1	S44092	probable carrier p	
8	95	13.7	296	2	B96830	hypothetical prote	
9	95	13.7	307	2	B60949	probable phosphate	
10	94.5	13.6	313	2	D84613	hypothetical prote	
11	94.5	13.6	351	2	T43493	hypothetical prote	
12	93	13.4	436	2	J01459	Btl protein precur	
13	90	12.9	311	2	G86383	probable mitochond	
14	90	12.9	363	2	T49281	mitochondrial phos	
15	89	12.8	310	2	T20229	hypothetical prote	
16	89	12.8	348	2	D84798	probable mitochond	
17	86.5	12.4	322	2	S57116	probable carrier p	
18	84	12.1	296	2	T23170	hypothetical prote	
19	83.5	12.0	353	2	T51393	probable mitochond	
20	83	11.9	161	1	AFKKB	allophycocyanin be	
21	83	11.9	284	2	T18253	probable mitochond	
22	83	11.9	288	2	T51595	phosphate transpor	
23	82.5	11.9	161	2	B44462	allophycocyanin be	
24	82.5	11.9	161	2	S33624	allophycocyanin be	
25	82	11.8	392	2	T05350	adenylate transloc	
26	80.5	11.6	162	2	C31305	allophycocyanin 1	
27	80.5	11.6	314	2	A36305	2-oxoglutarate/mal	
28	80	11.5	371	2	T05159	mitochondrial carr	
29	79.5	11.4	161	1	AFWMB	allophycocyanin be	

C. Accession: 800203
R.; Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141.MUID:21016719

Db 72 KNMLSQISHISTEGTLALWKGVSVILGAGPAHAHVFGTYEFCCKNLIDSSDTQTHHPF 131
QY 73 ANGAAGCVATLLHDAAMNPAEG-----NDSST-----YHVSVCSTC 108
Db 132 KTAISGACATTASDALMNPFDTIKORIQNLNTSASVWQTTKQIYQSEGLAAFYSTPTTLV 191
QY 109 ISLOFA-----EESTSVL 121
Db 192 MNIPFAAFNFVIESSTKFL 211

RESULT 5
T39149
Probable RNA splicing protein mitochondrial carrier protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39149
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: 221748
A:Accession: T39149
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-303 <OLI>
A:Cross-references: EMBL:Z99168; PIDN:CAB16300.1; GSPDB:GN00066; SPDB:SPAC8C9.12c
A:Experimental source: strain 972h-; cosmid c8C9
C:Genetics:
A:Gene: SPDB:SPAC8C9.12c
A:Map position: 1
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 17.1%; Score 119; DB 2; Length 303;
Best Local Similarity 27.6%; Pred. No. 0.00022;
Matches 34; Conservative 21; Mismatches 48; Indels 20; Gaps 3;
QY 14 NVLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKTLSDVIHPGGNSHIA 73
Db 60 NIVNSVIKISTEGVYSLRWGISSVIMGAGPSHAIFYSVLEPKSK-----INASPRPLA 115
QY 74 NGAAGCVATLLHDAAMNPAE-----GNDSSYHVSVCSTCISLQPAESTSVLGVNSVT 127
Db 116 SALAGACATITSDAFMTFVDIKQRMQLPSRKYKSALHC-----ATTVFRNEGLG 165
QY 128 LFY 130
Db 166 AFY 168

RESULT 6
S13533
mRNA splice defect-suppressing mitochondrial carrier MRS4 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKR052c
C:Species: Saccharomyces cerevisiae
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jul-2000
C:Accession: S13533; S38126
R:Wiesenberg, G.; Link, T.A.; von Ahsen, U.; Waldherr, M.; Schweyen, R.J.
J. Mol. Biol. 217, 23-37, 1991
A:Title: MRS4 and MRS4, two suppressors of mRNA splicing defects in yeast, are new members of the ADP/ATP carrier family
A:Reference number: S13532; MUID:91108815
A:Accession: S13533
A:Molecule type: DNA
A:Residues: 1-304 <IMO>
A:Cross-references: EMBL:X56444; NID:g3995; PIDN:CAA39828.1; PID:g3996
R:Visiers, S.; Urrestazu, L.A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38118
A:Accession: S38126
A:Molecule type: DNA
A:Residues: 1-304 <VIS>
A:Cross-references: EMBL:Z28277; NID:g486506; PID:g486507; MIPS:YKR052c
A:Experimental source: strain S288C
C:Genetics:

A:Gene: SGD:MRS4
A:Cross-references: SGD:S0001760; MIPS:YKR052c
A:Map position: 11R
A:Genome: nuclear
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: duplication; mitochondrion; transmembrane protein
F:20-109/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:117-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:206-301/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 15.3%; Score 106; DB 2; Length 304;
Best Local Similarity 32.5%; Pred. No. 0.0046;
Matches 25; Conservative 12; Mismatches 40; Indels 0; Gaps 0;
QY 15 VLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKTLSDVIHPGNSHIAN 74
Db 64 MTSQISKISTMEGSMALWGVQSVILGAGPAHAHVFGTYEFCCKARLISPEDMTHQPMKT 123
QY 75 GAAGCVATLLHDAAMNP 91
Db 124 ALSGTATTAAADALMNP 140

RESULT 7
S44092
Probable carrier protein c2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Oct-1999
C:Accession: S44092; T24753
R:Runswick, M.J.; Philippides, A.; Lauria, G.; Walker, J.E.
submitted to the EMBL Data Library, November 1993
A:Description: Extension of the mitochondrial transport superfamily: sequences of five members of the ADP/ATP carrier family
A:Reference number: S44090
A:Accession: S44092
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <RUN>
A:Cross-references: EMBL:X76116; NID:g472899; PIDN:CAA53722.1; PID:g472900
R:Lloyd, C.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z19932
A:Accession: T24753
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-384 <WIL>
A:Cross-references: EMBL:Z49070; PIDN:CAA88869.1; GSPDB:GN00020; CESP:T09F3.2
A:Experimental source: clone T09F3
C:Genetics:
A:Gene: CESP:T09F3.2
A:Map position: 2
A:Introns: 82/1, 113/1, 153/2, 179/2; 259/3; 359/3
C:Superfamily: Caenorhabditis probable carrier protein c2; ADP,ATP carrier protein repeat homology <ACP1>
F:4-43,141-193/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:45-67/Region: serine-rich
F:201-285/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:296-381/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 14.6%; Score 101.5; DB 1; Length 384;
Best Local Similarity 35.1%; Pred. No. 0.017;
Matches 33; Conservative 17; Mismatches 31; Indels 13; Gaps 5;
QY 5 QPDAARYRN-VLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLK--TLS 61
Db 137 QPDAARRGITVIKTYIKTEGIGALYKGLIPNLVGVAPSKAVFYTYTSTSKRFWES 196
QY 62 DVTHPGNSHIAN----GAAGCVATLLHDAAMNP 91
Db 197 EVLIP--NSAIVHWSAGSAGFVAA-----SAVNP 224
RESULT 8

```
Db      161 YKSVLDAITQWIRGEGVTSLRGRSSLTINRAMLVTSQLASYDSVKETILEKLLDKDLG 220
Qy      70 SHI-ANGAAGCVATLLHDAAMNPAE 93
       :: : ||| ||:: || :
Db      221 THVSASFAAGFVASV----ASNPD 241

RESULT 11
T43493
hypothetical protein DKFP243C119.1 - human
C:Species: Homo sapiens (man)
C:date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T43493
```


T20229
hypothetical protein C54G10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T20229
R:Matthews, L.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19240
A:Accession: T20229
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-310 <WIL>
A:Cross-references: EMBL:Z75532; PIDN:CAA99811.1; GSPDB:GN00023; CESP:C54G10.4
A:Experimental source: clone C54G10
C:Genetics:
A:Gene: CESP:C54G10.4
A:Map position: 5
A:Introns: 14/1; 63/1; 155/2; 256/2; 299/3
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 12.88; Score 89; DB 2; Length 310;
Best Local Similarity 25.58; Pred. No. 0.24; 55; Indels 38; Gaps 8;
Matches 39; Conservative 21; Mismatches 55; Indels 38; Gaps 8;

Qy 4 LQPDPA-ARYENVLEALWRIIRTEGLWRPMPRGLNVTATGAGPAHALYFACYEKLKK---- 58
:| | | : : : : :| | | :| | | : :
Db 125 IQDDKAHTKFNPIDATKQLLRTHGLKSLTRGFLATVARDAPAFGVYFASYENWARSCK 184
:| | | : : : : :| | | :| | | : :
Qy 59 -----TSLDVIHPGNGNSHIANGAGCVATLLH---DAAMNPAEGNDS--STYHSVGSCTC 108
||| | | | : : : : :| | | :| | | : :
Db 185 DGETSTLS-----SQQLIFAGGTAGMLSLWLFNYQTDIIKSRFQADNSYKSYM-----C 233
Qy 109 ISLOFAEES-----TSVLV-----GNSVTLF 129
| | | | | : | : | : | : | : | :
Db 234 IKQYILRGYRGFFVGLNSALIRAFPSNAATFF 266

Search completed: August 27, 2002, 03:57:45
Job time: 4951 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:16:49 ; Search time 37.5 Seconds
(without alignments)
135.260 Million cell updates/sec

Title: US-09-870-113-12

Perfect score: 695

Sequence: 1 MQSLQDPDPAARYNVLEALW.....QFAESTSVLVGNSVTLFYH 131

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	125.5	18.1	314	1 MRS3_YEAST	P10566 saccharomyc
2	106	15.3	304	1 MRS4_YEAST	P23500 saccharomyc
3	95	13.7	307	1 ODC2_YEAST	Q99297 saccharomyc
4	94.5	13.6	359	1 CG69_HUMAN	Q9b214 homo sapien
5	93	13.4	436	1 BTL1_MAIZE	P29518 zea mays (m
6	90	12.9	320	1 DNC_HUMAN	Q9hc21 homo sapien
7	89.5	12.9	161	1 PHAB_CYACA	Q9tl58 cyanidium c
8	86.5	12.4	161	1 PHAB_GALSU	P00319 galdieria s
9	86.5	12.4	322	1 SFC1_YEAST	P33303 saccharomyc
10	82.5	11.9	161	1 PHAB_SYNY3	Q01952 synchocyst
11	82.5	11.9	161	1 PHAB_SYNY4	Q02924 synchocyst
12	80.5	11.6	162	1 PHAB_FREDI	P16571 fremyella d
13	80.5	11.6	313	1 M2OM_BOVIN	P22292 bos taurus
14	79.5	11.4	161	1 PHAB_MASLA	P00318 mastigoclad
15	79.5	11.4	313	1 M2OM_HUMAN	Q02978 homo sapien
16	79.5	11.4	328	1 Y051_CAPEL	Q09461 caenorhabdi
17	79	11.4	317	1 GGH_RAT	Q62867 rattus norv
18	78.5	11.3	161	1 PHAB_CYPAP	P00320 cyanophora
19	78	11.2	366	1 YG5F_YEAST	P53320 saccharomyc
20	77.5	11.2	301	1 ADT_ANOGA	Q27238 anopheles g
21	77.5	11.2	313	1 M2OM_MOUSE	Q9cr62 mus musculu
22	77	11.1	368	1 YM39_YEAST	Q03829 saccharomyc
23	76.5	11.0	161	1 PHAB_SYNEL	P50031 synchococc
24	76.5	11.0	1320	1 PUTA_SALTY	P10503 salmonella
25	76	10.9	315	1 SA18_HUMAN	Q9h1k4 homo sapien
26	75.5	10.9	161	1 PHAB_AGLNE	P28556 aglaothamni
27	75.5	10.9	161	1 PHAB_ANASP	P80557 anabaena sp
28	75.5	10.9	161	1 PHAB_ANAVA	P00317 anabaena va
29	75.5	10.9	161	1 PHAB_SYNP6	P06113 synchococc
30	74.5	10.7	161	1 PHAB_SPLPL	P72505 spirulina p
31	74	10.6	309	1 UCP2_PIG	Q97562 sus scrofa
32	73.5	10.6	161	1 PHAB_ANACY	P07326 anabaena cy
33	73.5	10.6	287	1 DIC_HUMAN	Q9ubx3 homo sapien

34	73.5	10.6	287	1 DIC_MOUSE	Q9qzd8 mus musculu
35	73.5	10.6	313	1 M2OM_RAT	P97700 rattus norv
36	73	10.5	312	1 SA18_MOUSE	Q9db41 mus musculu
37	72	10.4	695	1 CMCI_DROME	Q9va71 drosophila
38	71.5	10.3	315	1 MFT_HUMAN	Q9h2d1 homo sapien
39	71	10.2	308	1 UCP3_RAT	P56499 rattus norv
40	70	10.1	299	1 YFBH_SALTY	O52326 salmonella
41	70	10.1	301	1 FPG_RHIME	O59752 rhizobium m
42	70	10.1	308	1 ADT_CHLRE	P27080 chlamydomon
43	69.5	10.0	292	1 ORT1_YEAST	Q12375 saccharomyc
44	69.5	10.0	299	1 TXTP_YEAST	P38152 saccharomyc
45	69.5	10.0	1312	1 PUTA_KLEAE	O52485 klebsiella

ALIGNMENTS

RESULT 1
MRS3_YEAST
ID MRS3_YEAST STANDARD; PRT; 314 AA.
AC P10566;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Mitochondrial RNA splicing protein MRS3.
GN MRS3 OR YJL133W OR J0675.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IC8/R101;
RX MEDLINE=88121698; PubMed=2448588;
RA Schmidt C., Soellner T., Schweyen R.J.;
RT "Nuclear suppression of a mitochondrial RNA splice defect: nucleotide
sequence and disruption of the MRS3 gene.";
RL Mol. Gen. Genet. 210:145-152(1987).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RC STRAIN=MI301;
RX MEDLINE=91108815; PubMed=1703236;
RA Wiesenberger G., Link T.A., von Ahsen U., Waidherr M., Schweyen R.J.;
RT "MRS3 and MRS4, two suppressors of mtRNA splicing defects in yeast,
are new members of the mitochondrial carrier family.";
RL J. Mol. Biol. 217:23-37(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=96408771; PubMed=8813765;
RA Katsoulou C., Tzermia M., Tavernarakis N., Alexandraki D.;
RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast
chromosome X reveals 14 known genes and 13 new open reading frames
including homologues of genes clustered on the right arm of
chromosome XI.";
RL Yeast 12:787-797(1996).
CC -1- FUNCTION: MRS3 SUPPRESSES A MITOCHONDRIAL SPLICE DEFECT IN THE
FIRST INTRON OF THE COB GENE. IT MAY ACT AS A CARRIER. EXERTING
ITS SUPPRESSOR ACTIVITY VIA MODULATION OF SOLUTE CONCENTRATIONS IN
THE MITOCHONDRIUM (POSSIBLY OF CATIONS).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,
TO YEAST MRS4; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER
PROTEIN.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch).

```
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT VARSPLOC 64 72 LPSLOSTG -> W (IN ISOFORM 2).
FT VARIANT 247 L -> F.
FT CONFLICT 266 266 /FTID=VAR_012756.
FT SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;

Query Match 13.6%; Score 94.5; DB 1; Length 359;
Best Local Similarity 29.6%; Pred. No. 0.012;
Matches 34; Conservative 15; Mismatches 39; Indels 27; Gaps 6;

QY 7 DPAARYNVLEALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL--- 60
DB 100 DP-TFTGTMDAFYKIVRHEGTRTLW---SGLPATLVMTVPATAIYFAYDQLKFLCGR 155
QY 61 ---SDVTHPGGSHIANGAAGCVATLLHDAAMNPAE-----GNDSTSYHSVSGSC 106
DB 156 ALTSOLIAP-----MWAGALARLGTIVVISPLELMTKQLQAHVSYRELGCAC 202

RESULT 5
BT1_MAIZE
ID B11_MAIZE STANDARD; PRT; 436 AA.
AC P29518;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brittle-1 protein, chloroplast precursor.
GN Brl.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005685; PubMed=1668652;
RA Sullivan T.D., Strelow L.I., Illingworth C.A., Phillips R.L.,
RA Nelson O.E. Jr.,
RT "Analysis of maize brittle-1 alleles and a defective
RT Suppressor-mutator-induced mutable allele.";
RL Plant Cell 3:1337-1348(1991).
CC -1- FUNCTION: COULD PLAY A ROLE IN AMYLOPLAST MEMBRANE TRANSPORT.
CC -1- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
CC -1- TISSUE SPECIFICITY: ENDOSPERM OF DEVELOPING KERNELS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M79333; AAA33438.1; -.
CC DR PIR; JQ1459; JQ1459.
CC DR MaizeDB; 47578; -.
CC DR InterPro; IPR002067; Mit_carrier.
CC DR Pfam; PF00153; mito_carr; 3.
CC DR PRINTS; PR00926; MITOCARRIER.
CC DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Transit peptide; Chloroplast; Amyloplast; Transmembrane.
FT TRANSIT 1 75 CHLOROPLAST (POTENTIAL).
FT CHAIN 76 436 BRITTLE-1 PROTEIN.
FT TRANSMEM 229 247 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
```

```
SQ SEQUENCE 436 AA; 46627 MW; 9600C05F603E9DAE CRC64;

Query Match 13.4%; Score 93; DB 1; Length 436;
Best Local Similarity 41.2%; Pred. No. 0.021;
Matches 21; Conservative 7; Mismatches 23; Indels 0; Gaps 0;

QY 12 YRNVEALWRIIRTEGLRPMRGLNVTATGAGPAHALYFACYEKLKKTLSLSD 62
DB 365 YQNVLHAIYCLIKKEGAGGLYRGIGSPGSCIKLMPAAGIAFWCYEACKKILVD 415

RESULT 6
DNC_HUMAN
ID DNC_HUMAN STANDARD; PRT; 320 AA.
AC Q9HC21;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial deoxynucleotide carrier (Mitochondrial uncoupling
DE protein 1).
GN SLC25A19 OR DNC OR MUP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RX TISSUE=Liver;
RX PubMed=11226231;
RA Dolce V., Fiermonte G., Runswick M.J., Palmieri F., Walker J.E.;
RT "The human mitochondrial deoxynucleotide carrier and its role in the
RT toxicity of nucleoside antivirals.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:2284-2288(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Renard S., Mondesert G., Besnard F.;
RT "MUP 1, a mitochondrial uncoupling protein.";
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Responsible for the uptake of deoxynucleotides into the
CC matrix of the mitochondria. Transports all four deoxy NDPs, and,
CC less efficiently, the corresponding dNTPs. Does not transport
CC dNMPs, NMPs, deoxynucleosides, nucleosides, purines, or
CC pyrimidines. Supply deoxynucleotides to the mitochondrial matrix
CC for conversion to triphosphates and incorporation into
CC mitochondrial DNA.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- TISSUE SPECIFICITY: Expressed in all tissues examined except for
CC placenta. Highest levels in colon, kidney, lung, testis, spleen,
CC and brain.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- DISEASE: Likely to be medically important by providing the means
CC of uptake into mitochondria of nucleoside analogs, leading to the
CC mitochondrial impairment that underlies the toxic side effects of
CC such drugs in the treatment of viral illnesses, including AIDS,
CC and in cancer therapy.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ251857; CAC27560.1; -.
CC DR
```



```
[2]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=97061201; PubMed=8905231;
RA  Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA  Miyajima N., Hikosawa M., Sugliura M., Sasamoto S., Kimura T.,
RA  Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA  Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA  Yamada M., Yasuda M., Tabata S.;
RT  "Sequence analysis of the genome of the unicellular cyanobacterium
RT  Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT  entire genome and assignment of potential protein-coding regions.";
RL  DNA Res. 3:109-136(1996).
[3]
RN  SEQUENCE OF 1-20.
RP  MEDLINE=97443974; PubMed=9298645;
RA  Sazuka T., Ohara O.;
RT  "Towards a proteome project of cyanobacterium Synechocystis sp.
RT  strain PCC6803: linking 130 protein spots with their respective
RT  genes.";
RL  Electrophoresis 18:1252-1258(1997).
CC  -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC  FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC  ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
CC  -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC  -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M71135; AAA27277.1; -
DR  EMBL; D90910; BAA17875.1; -
DR  PIR; B44462; B44462.
DR  HSSP; P00318; 1B33.
DR  InterPro; IPR001659; Phycobilisome.
DR  Pfam; PF00502; Phycobilisome; 1.
DR  Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW  Methylation.
FT  MOD_RES 71 71 METHYLATION (BY SIMILARITY).
FT  BINDING 81 81 PHYCOCYANOBILIN CHROMOPHORE
FT  (BY SIMILARITY).
SQ  SEQUENCE 161 AA; 17215 MW; 63F01E5903BA1B83 CRC64;
-----
Query Match 11.9%; Score 82.5; DB 1; Length 161;
Best Local Similarity 29.0%; Pred. No. 0.085;
Matches 27; Conservative 15; Mismatches 24; Indels 27; Gaps 4;
QY 54 EKLKTL--SDVIHPGGNSHTANGAGCV-----ATLLHDAAMNPAE 93
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54 EAVAKSLYSVDTRPGGNYTTRRYAACIRDLDDYLYRVATYAMLADASILDRLVINGLK 113
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 94 GNDSTYHSVG---SCTCISLQFAEESTSVLVG 123
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 114 ----ETYNLSGLVPSISSTVQAIKEVTASLVG 142
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 12
PHAB_SYNY4 STANDARD; PRT; 161 AA.
AC Q02924;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Allophycocyanin beta chain.
GN APCB.
OS Synechocystis sp. (strain PCC 6714).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1147;
```

```
[1]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=93222481; PubMed=8467079;
RA  Dinagno L.M., Haselkorn R.;
RT  "Isolation and characterization of the genes encoding allophycocyanin
RT  subunits and two linker proteins from Synechocystis 6714.";
RL  Plant Mol. Biol. 21:835-846(1993).
CC  -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC  FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC  ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
CC  -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC  -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L02308; AAA69683.1; -
DR  HSSP; P00318; 1B33.
DR  InterPro; IPR001659; Phycobilisome.
DR  Pfam; PF00502; Phycobilisome; 1.
DR  Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW  Methylation.
FT  MOD_RES 71 71 METHYLATION (BY SIMILARITY).
FT  BINDING 81 81 PHYCOCYANOBILIN CHROMOPHORE.
FT  (BY SIMILARITY).
SQ  SEQUENCE 161 AA; 17242 MW; 63F0047008630A73 CRC64;
-----
Query Match 11.9%; Score 82.5; DB 1; Length 161;
Best Local Similarity 29.0%; Pred. No. 0.085;
Matches 27; Conservative 15; Mismatches 24; Indels 27; Gaps 4;
QY 54 EKLKTL--SDVIHPGGNSHTANGAGCV-----ATLLHDAAMNPAE 93
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54 EAVAKSLYSVDTRPGGNYTTRRYAACIRDLDDYLYRVATYAMLADASILDRLVINGLK 113
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 94 GNDSTYHSVG---SCTCISLQFAEESTSVLVG 123
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 114 ----ETYNLSGLVPSISSTVQAIKEVTASLVG 142
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 12
PHAB_FREDI STANDARD; PRT; 162 AA.
AC P16571;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Allophycocyanin beta chain.
GN APCB1.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Fremyella.
OX NCBI_TaxID=11197;
RN  SEQUENCE FROM N.A.
RP  MEDLINE=89053869; PubMed=2461358;
RA  Hounard J., Capuano V., Coursin T., Tandeau de Marsac N.;
RT  "Genes encoding core components of the phycobilisome in the
RT  cyanobacterium Calothrix sp. strain PCC 7601: occurrence of a
RT  multigene family.";
RL  J. Bacteriol. 170:5512-5521(1988).
CC  -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC  FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC  ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
CC  -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC  -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
```


SEQUENCE OF 55-82, AND METHYLATION.
RA Ruemelli R., Suter F., Wirth M., Sidler W., Zuber H.;
RT "Gamma-N-methylasparagine in phycobiliproteins from the cyanobacteria
RT Mastigocladus laminosus and Calothrix.";
RL FEBS Lett. 221:1-2(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RC STRAIN=PCC 7603;
RX MEDLINE=99145555; PubMed=9990029;
RA Reuter W., Wiegand G., Huber R., Than M.E.;
RT "Structural analysis at 2.2 A of orthorhombic crystals presents the
RT asymmetry of the allophycocyanin-linker complex, AP-1C7.8, from
RT phycobilisomes of Mastigocladus laminosus.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1363-1368(1999).
CC FROM THE PHYCOCYANIN-PROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
DR PIR: S02501; AFMBW.
DR PIR: S02612; S02612.
DR PDB: 1B33; 23-FEB-99.
DR InterPro: IPR001659; Phycobilisome.
DR Pfam: PF00502; Phycobilisome; 1.
KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW Methylation; 3D-structure.
FT MOD_RES 71 71 METHYLATION.
FT BINDING 81 81 PHYCOCYANOBILIN CHROMOPHORE.
FT CONFLICT 59 60 SL -> KT (IN REF. 1).
FT CONFLICT 67 67 R -> L (IN REF. 1).
FT CONFLICT 71 71 N -> D (IN REF. 1).
SQ SEQUENCE 161 AA; 17374 MW; 18E7DD78D7178350 CRC64;

Query Match 11.4%; Score 79.5; DB 1; Length 161;
Best Local Similarity 31.5%; Pred. No. 0.18;
Matches 28; Conservative 12; Mismatches 30; Indels 19; Gaps 4;

QY 54 EKLKKTLL--SDVTHPGNGSHIANGAAGCVATL---LHDAAMNPAGNDS----- 97
DB 54 EAVAKSLYSDFTRPGNMYTTRRYAACIRDLDYLYRYATYAMLAGDPSILDERVLNGLK 113
QY 98 SYTHSVG---SCTCISLQFAESTSVLVG 123
DB 114 ETYNSLGVPISTVQAIQAKVETASLVG 142

RESULT 15
MZOM_HUMAN
ID M2OM_HUMAN STANDARD; PRT; 313 AA.
AC Q02978; 075537;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial 2-oxoglutarate/malate carrier protein (OGCP).
GN SLIC25A11 OR SLIC20A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93091249; PubMed=1457818;
RA Iacobazzi V., Palmieri F., Runswick M.J., Walker J.E.;
RT "Sequences of the human and bovine genes for the mitochondrial 2-
RT oxoglutarate carrier.";
RL DNA Seq. 3:79-88(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Yu W., Gibbs R.A.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE TRANSPORT OF 2-OXOGLUTARATE ACROSS THE

INNER MITOCHONDRIAL MEMBRANE IN AN ELECTRONEUTRAL EXCHANGE FOR
MALATE OR OTHER DICARBOXYLIC ACIDS, AND PLAYS AN IMPORTANT ROLE
IN SEVERAL METABOLIC PROCESSES, INCLUDING THE MALATE-ASPARTATE
SHUTTLE, THE OXOGLUTARATE/ISOCITRATE SHUTTLE, IN GLUCONEOGENESIS
FROM LACTATE, AND IN NITROGEN METABOLISM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66114; CAA46905.1; -.
DR EMBL: AF070548; AAC28637.1; -.
DR PIR: S29598; S29598.
DR MIM: 604165; -
DR InterPro: IPR001993; Mitoch_carr.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; 2.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT INIT_MET 0 0 BY SIMILARITY.
FT TRANSMEM 23 41 POTENTIAL.
FT TRANSMEM 82 100 POTENTIAL.
FT TRANSMEM 118 139 POTENTIAL.
FT TRANSMEM 182 201 POTENTIAL.
FT TRANSMEM 221 239 POTENTIAL.
FT TRANSMEM 280 299 POTENTIAL.
FT REPEAT 1 107 1.
FT REPEAT 108 208 2.
FT REPEAT 209 308 3.
FT CONFLICT 11 11 I -> M (IN REF. 1).
SQ SEQUENCE 313 AA; 33930 MW; A4831A2E1A9F175A CRC64;

Query Match 11.4%; Score 79.5; DB 1; Length 313;
Best Local Similarity 31.1%; Pred. No. 0.39;
Matches 28; Conservative 12; Mismatches 49; Indels 1; Gaps 1;

QY 4 LQDPDPAARYNNVLEALWRIIRTEGLWRPMRGLNVTATGAPAHALYFACYEKIKKTLSDV 63
DB 152 LPADQRRGKYNFVNALIRITREEGVLTLMRGCIPTMARAVVWVNAQAQLASYSQSKQFLDS 211
QY 64 IHPGGNSHIANGAAGCVATLLHDAAMNPAE 93
DB 212 GYFSDNI-LCHFCAISMISGLVTTAASMPVD 240

Search completed: August 27, 2002, 04:16:49
Job time: 1241 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:19:24 ; Search time 130.92 seconds
(without alignments)
173.101 Million cell updates/sec

Title: US-09-870-113-12
Perfect score: 695
Sequence: 1 MQSLQPDPAARYRNVEALW.....QFAEESTVLGVNSVTLEFVH 131

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues 562222
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriaph.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	505	72.7	364	4 Q96A46	Q96A46 homo sapien
2	351	50.5	347	4 Q9NYZ2	Q9NYZ2 homo sapien
3	348	50.1	338	11 Q920G8	Q920G8 mus musculus
4	280	40.3	155	4 Q969S1	Q969S1 homo sapien
5	276	39.7	182	11 Q912Y0	Q912Y0 mus musculus
6	216.5	31.2	312	5 Q23125	Q23125 caenorhabdi
7	178.5	25.7	379	5 Q9VAY3	Q9VAY3 drosophila
8	169	24.3	380	5 Q9NYH6	Q9NYH6 onchocerca
9	168	24.2	303	5 Q94638	Q94638 onchocerca
10	166	23.9	301	5 Q94634	Q94634 onchocerca
11	143.5	20.6	781	10 Q9LMJ6	Q9LMJ6 arabidopsis
12	141.5	20.4	289	10 O82049	O82049 ribes nigru
13	139	20.0	331	10 O64731	O64731 arabidopsis
14	119	17.1	303	3 O14281	O14281 schizosacch
15	115	16.5	320	11 Q921P8	Q921P8 mus musculus
16	114	16.4	345	11 Q9CYJ1	Q9CYJ1 mus musculus

17	108	15.5	321	4 Q9BSK2	Q9BSK2 homo sapien
18	101.5	14.6	384	5 Q27244	Q27244 caenorhabdi
19	99.5	14.3	322	10 Q9ARL9	Q9ARL9 horddeum vul
20	99.5	14.3	413	5 Q9VWF9	Q9VWF9 drosophila
21	99	14.2	307	4 O14589	O14589 homo sapien
22	98	14.1	322	10 Q9AX03	Q9AX03 oryza sativ
23	96	13.8	338	4 Q9UHR1	Q9UHR1 homo sapien
24	95.5	13.7	310	4 Q9NVN5	Q9NVN5 homo sapien
25	95.5	13.7	311	4 Q96CQ1	Q96CQ1 homo sapien
26	95.5	13.7	311	11 Q922G0	Q922G0 mus musculu
27	95	13.7	296	10 Q9CA93	Q9CA93 arabidopsis
28	94.5	13.6	313	10 Q9SJY5	Q9SJY5 arabidopsis
29	94.5	13.6	313	10 Q94K32	Q94K32 arabidopsis
30	94.5	13.6	337	4 Q9P182	Q9P182 homo sapien
31	94.5	13.6	351	4 Q9Y379	Q9Y379 homo sapien
32	94.5	13.6	351	4 Q9UF66	Q9UF66 homo sapien
33	94.5	13.6	359	4 Q9BZJ4	Q9BZJ4 homo sapien
34	94	13.5	318	11 Q9DAM5	Q9DAM5 mus musculu
35	93.5	13.5	336	11 Q9D981	Q9D981 mus musculu
36	91	13.1	132	4 Q9H0G8	Q9H0G8 homo sapien
37	90	12.9	311	10 Q9C6K8	Q9C6K8 arabidopsis
38	90	12.9	320	4 Q9HC21	Q9HC21 homo sapien
39	90	12.9	363	10 Q9M2Z8	Q9M2Z8 arabidopsis
40	90	12.9	376	10 Q9FR17	Q9FR17 arabidopsis
41	89.5	12.9	359	11 Q9D8K8	Q9D8K8 mus musculu
42	89	12.8	310	5 O18844	O18844 caenorhabdi
43	89	12.8	348	10 Q9SH98	Q9SH98 arabidopsis
44	89	12.8	358	5 Q9VQ37	Q9VQ37 drosophila
45	89	12.8	365	5 Q95TJ5	Q95TJ5 drosophila

ALIGNMENTS

RESULT 1

Q96A46 ID Q96A46 PRELIMINARY; PRT; 364 AA.
AC Q96A46;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE CARRIER SPLICER VARIANT).
GN HMRS3/4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21195335; PubMed=11297739;
RA Li F., Nikali K., Grogan J., Leibiger I., Leibiger B., Schweyen R.,
LA Larsson C., Suomalainen A.;
RT "Characterization of a novel human putative mitochondrial transporter homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.";
RL FEBS Lett. 494:79-84(2001).
DR ENBL; A330307; CAC27996.1;
DR ENBL; AF327402; AAK49519.1;
SQ SEQUENCE 364 AA; 39271 MW; 22049C4F8DD7A577 CRC64;

Query Match 72.7%; Score 505; DB 4; Length 364;
Best Local Similarity 87.5%; Pred. No. 5.3e-45;
Matches 99; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

Qy 1 MQSLQPDPAARYRNVEALWRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
|||||
Db 100 MQSLQPDPAARYRNVEALWRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 159
|||||
Qy 61 SDVTHPGGNSHIANGAGCVATLLHDAAMPAAE-----GNDSTYHSVSC 106
|||||
Db 160 SDVTHPGGNSHIANGAGCVATLLHDAAMPAAEYVVKORMQMYNSPYHRVTDC 211
|||||

```

RESULT 2
Q9NYZ2
ID Q9NYZ2; PRELIMINARY; PRT; 347 AA.
AC Q9NYZ2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HT015 PROTEIN.
GN HT015.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HYPOTHALAMUS;
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human hypothalamus.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223466; AAF64141.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 347 AA; 37828 MW; 700DE61B230E001E CRC64;

Query Match 50.5%; Score 351; DB 4; Length 347;
Best Local Similarity 67.7%; Pred. No. 8.2e-29;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGVLNVTATGAGPAHALYFACYEKLKRTL 60
Db 73 MOSLSPDKAQYTSYGALKIMRTGEFWRPLRGVNVIMGAGPAHAMFYFACYENKRTL 132

QY 61 SDVIHPGNSHTANGACVATLLHDAAMNPAE 93
Db 133 NDVFFHQGNHSLANGVAGSMATLLHDAAMNPAE 165

RESULT 3
Q920G8
ID Q920G8; PRELIMINARY; PRT; 338 AA.
AC Q920G8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Q.-Z., Ruan Q.-G., Eckenrode S., Shi J.-D., Cruz P., Wang C.-Y.,
RA She J.-X.;
RT "A new gene which is highly expressed in NOD mice spleen.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF288621; AAL23859.1; -.
SQ SEQUENCE 338 AA; 37510 MW; BB35B1F70C56A3FE CRC64;

Query Match 50.1%; Score 348; DB 11; Length 338;
Best Local Similarity 68.8%; Pred. No. 1.6e-28;
Matches 64; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGVLNVTATGAGPAHALYFACYEKLKRTL 60
Db 73 MOSLNPDPKARYTSYGALKRIMHTEGFWRPLRGVLNVMVMGAGPAHAMFYFACYENKRTL 132

QY 61 SDVIHPGNSHTANGACVATLLHDAAMNPAE 93
Db 133 NDVFFHQGNHSLANGVAGSMATLLHDAAMNPAE 165

```

```

RESULT 4
Q969S1
ID Q969S1; PRELIMINARY; PRT; 155 AA.
AC Q969S1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MITOCHONDRIA SOLUTE CARRIER PROTEIN (HYPOTHETICAL 16.8 KDA
DE PROTEIN).
GN MSCP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Q., Eckenrode S., Ruan Q., Wang C., Shi J., McIndoe R.A., She J.;
RT "Molecular cloning of a novel mitochondria solute carrier protein
RT (MSCP) gene from mouse and human and its down-regulation in mouse
RT spleen during the maturation of the immune system.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS, AND LEIOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY032628; AAK38154.1; -.
DR EMBL; BC015013; AAH15013.1; -.
KW Hypothetical protein.
SQ SEQUENCE 155 AA; 16832 MW; 8ACB98A483C8E6EF CRC64;

Query Match 40.3%; Score 280; DB 4; Length 155;
Best Local Similarity 65.3%; Pred. No. 9.2e-22;
Matches 49; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGVLNVTATGAGPAHALYFACYEKLKRTL 60
Db 73 MOSLSPDKAQYTSYGALKIMRTGEFWRPLRGVNVIMGAGPAHAMFYFACYENKRTL 132

QY 61 SDVIHPGNSHTANG 75
Db 133 NDVFFHQGNHSLANG 147

RESULT 5
Q912Y0
ID Q912Y0; PRELIMINARY; PRT; 182 AA.
AC Q912Y0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MITOCHONDRIAL CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Li Q., Eckenrode S., Wang C., Ruan Q., Shi J., McIndoe R.A., She J.;
RT "A novel mouse mitochondrial carrier protein gene is up-regulated from
RT young to adult NOD mice.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF361699; AAL27990.1; -.
SQ SEQUENCE 182 AA; 19838 MW; 18E2C5E801228693 CRC64;

Query Match 39.7%; Score 276; DB 11; Length 182;
Best Local Similarity 66.7%; Pred. No. 2.9e-21;
Matches 50; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

```

```

QY 1 MOSLOPDPARYRNVLALRIITTEGLWRPRLNVTATGAGPAHALYFACYEKLKTL 60
Db 73 MQSLNPPKARYSIYVGLAKRIMHTEGFWRLGLNMMGAGPAHALYFACYEMKRTL 132
QY 61 SDVHPGNGSHANG 75
Db 133 NDVFSHOGNSHLANG 147

RESULT 6
Q23125 PRELIMINARY; PRT; 312 AA.
AC Q23125;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE W02B12.9 (TREMBLrel. 19, Last annotation update)
GN W02B12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN RP
RA Swinburne J., Ainscough R.;
RL Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
[2]
RN RP
RA none;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z66521; CAAG1399.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;

Query Match 31.2%; Score 216.5; DB 5; Length 312;
Best Local Similarity 50.0%; Pred. No. 1e-14;
Matches 47; Conservative 16; Mismatches 28; Indels 3; Gaps 2;

QY 1 MOSLOPDPARYRNVLALRIITTEGLWRPRLNVTATGAGPAHALYFACYEKLKTL 60
Db 45 MOSLCPETKCPKPTVHSLMSIVKREGWRLRGVNAAGSMPAHALYFTVYKMGYL 104
QY 61 SDVHPGNGSH-ANGAAGCVATLLHDAANPAE 93
Db 105 TG--NSAGHSNTLAYGASGVVATLIHDAINPAE 136

RESULT 7
Q9VAY3 PRELIMINARY; PRT; 379 AA.
AC Q9VAY3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
GN CG4963.
DE CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN RP
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA

```

```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier K., Tector A.C., Turner R., Venter E., Wang A.H., Wang X.,
RA Svirskas R., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Wang Z.-Y., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pauleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; AF003763; AAF56764.1; -.
DR EMBL; AY060268; AAL25307.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCH_CARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 25.7%; Score 178.5; DB 5; Length 379;
Best Local Similarity 39.3%; Pred. No. 1.3e-10;
Matches 44; Conservative 17; Mismatches 40; Indels 11; Gaps 3;

QY 1 MOSLOPDPARYRNVLALRIITTEGLWRPRLNVTATGAGPAHALYFACYEKLKTL 60
Db 42 MQSL--SPTKNNMIVSTLTMTREGLLRPIRGASAVLGAAGPAHSLFAAEMTKELT 99
QY 61 SDVHPGNGSHANGAAGCVATLLHDAANPAE-----GNDSSTYHVSVC 106
Db 100 AKFTSVRNLYVISGA--VATLIHDAISSPTDVIKQRMQMYNSPYTSVSC 148

```

```

RESULT 8
Q9NHV6
ID Q9NHV6 PRELIMINARY; PRT; 380 AA.
AC Q9NHV6
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-FEB-1997 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 41.8 KDA PROTEIN.
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Schmid K.J., Tautz D.;
RT "A screen for rapidly evolving genes from Drosophila.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF217402; AAF73387.1; -
DR FlyBase: FBgn0039561; CG4963.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PRINTS: PR00926; MITOCARRIER.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 41844 MW; 4C9AA524B97F8C6C CRC64;

Query Match 24.3%; Score 169; DB 5; Length 380;
Best Local Similarity 38.1%; Pred. No. 1.3e-09;
Matches 43; Conservative 17; Mismatches 41; Indels 12; Gaps 4;

Qy 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGUNVTATGAGPAHALYFACYEKLKRTL 60
Db 42 MOSL--SPPTKNNMIVSLTRMITREGLRLPIRGASAVVLGAGPTSHLYFAAYETKLT 99
Qy 61 SDVIHPGNSHIANGAACGVATLLHDAAMNPAE-----GNDSSP--YHSVCS 106
Db 100 AKFTSVRLNLYVIGA---VATLHDAISSPTDIKTAAYADCTTRTPHPVWSC 149

RESULT 9
Q94638
ID Q94638 PRELIMINARY; PRT; 303 AA.
AC Q94638
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL Mitochondrial solute carriers.";
DR EMBL: U45998; AAB19037.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 303 AA; 33861 MW; 224BFA547E5D617C CRC64;

Query Match 24.2%; Score 168; DB 5; Length 303;
Best Local Similarity 36.6%; Pred. No. 1.3e-09;

Qy 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGUNVTATGAGPAHALYFACYEKLKRTL 60
Db 42 MOSL--SPPTKNNMIVSLTRMITREGLRLPIRGASAVVLGAGPTSHLYFAAYETKLT 99
Qy 61 SDVIHPGNSHIANGAACGVATLLHDAAMNPAE 93
Db 100 AKFTSVRLNLYVIGA---VATLHDAISSPTDIKTAAYADCTTRTPHPVWSC 149

RESULT 10
Q94634
ID Q94634 PRELIMINARY; PRT; 301 AA.
AC Q94634
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca gibsoni.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6284;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL Mitochondrial solute carriers.";
DR EMBL: U45997; AAB19036.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 301 AA; 34176 MW; C0BA8D819FB8EA79 CRC64;

Query Match 23.9%; Score 166; DB 5; Length 301;
Best Local Similarity 36.6%; Pred. No. 2.1e-09;
Matches 34; Conservative 20; Mismatches 37; Indels 2; Gaps 1;

Qy 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGUNVTATGAGPAHALYFACYEKLKRTL 60
Db 42 LQSLCPCPETSCPTAMHSLMSVKGREGLRLSLKGVNAVVLGTIPAHAFYTYVYENSKAYL 101
Qy 61 SDVIHPGNSHIANGAACGVATLLHDAAMNPAE 93
Db 102 LN--NPRVSNVSVAISGALATVIHDAVMNPAE 132

RESULT 11
Q9LMJ6
ID Q9LMJ6 PRELIMINARY; PRT; 781 AA.
AC Q9LMJ6
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F10K1.26 PROTEIN.
GN F10K1.26.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Liu S.X., Chan A., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
RA Sakano H., Toriumi M., Chung M., Goldsmith A., Liu A., Liu A.,
RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "The sequence of BAC F10K1 from Arabidopsis thaliana chromosome 1.";
```

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC067971; AAF82217.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR002067; Mit_carrier.
DR Pfam: PF00153; mito_carr; 4.
DR PRINTS: PR00926; MITOCARRIER.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 781 AA; 87081 MW; 9FB579B9BD746D1E CRC64;

Query Match 20.6%; Score 143.5; DB 10; Length 781;
Best Local Similarity 33.6%; Pred. No. 1.5e-06;
Matches 49; Conservative 16; Mismatches 52; Indels 29; Gaps 6;
QY 1 MQSLQDPDPAARYNVLEALWRIIRTEGLWPRMGLNVTATGAGPAHALYFACYEKLKTL 60
DB 64 MQALRCP-LKPVGIRAFKSIQKEGSAHYRGWGLGAGPAHAYFSFVSVSKYL 122
QY 61 SDVIHPGG--NSHTANGAACVATLLHDAAMNP-----AEGNDSSTYHSVGSCT- 107
DB 123 S-----AGDQNNVAHAMSGVFATISSDAVTPMDVMVKQLQMGEG---TYKGVWDCVK 173
QY 108 -----CISLQFAESTSVLGNVST 127
DB 174 RVLREGIGAFYASYRTTVMNAPFT 199

RESULT 12
082049 PRELIMINARY; PRT; 289 AA.
AC 082049;
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE MITOCHONDRIAL CARRIER PROTEIN.
GN Ribes nigrum.
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Saxifragales; Grossulariaceae; Ribes.
OX NCBI_TaxID=78511;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BEN ALDER; TISSUE=FRUIT;
RA Woodhead M.R.; Taylor M.A.; Brennan R.M.; McNicol R.J.; Davies H.V.;
RL Thesis (1995), University of Dundee, UK.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BEN ALDER; TISSUE=FRUIT;
RA Woodhead M.R.; Taylor M.A.; Brennan R.M.; McNicol R.J.; Davies H.V.;
RT "Cloning and characterisation of the cDNA clones of five genes that
RT are differentially expressed during ripening in the fruit of
RT blackcurrant (Ribes nigrum L.).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL: AJ007580; CAA07568.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 289 AA; 30383 MW; 5666EAB03DCC507C CRC64;

Query Match 20.4%; Score 141.5; DB 10; Length 289;
Best Local Similarity 31.2%; Pred. No. 7.5e-07;
Matches 40; Conservative 18; Mismatches 51; Indels 19; Gaps 3;
QY 9 AARYNVLEALWRIIRTEGLWPRMGLNVTATGAGPAHALYFACYEKLKTLSDVIHPGG 68
DB 32 SAQSAGLRQALGSLIKVEGPAGLYRGIGMGLGAGPAHAYFSVYEMKETFS---HGPD 88

QY 69 NSHTANGAACVATLLHDAAMNPAE-----GNDSSYHSVGSCTCISLQFAESTSVLV 122
DB 89 SNSCAHAVSGVFATVSDAVITPMDVMVKQLQLOSSPKGVDCV-----RRVLV 138
QY 123 GNSVTLFY 130
DB 139 EEGIGAFY 146
RESULT 13
064731 PRELIMINARY; PRT; 331 AA.
AC 064731;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE PUTATIVE MITOCHONDRIAL CARRIER PROTEIN.
GN AT2G30160.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X.; Kaul S.; Rounsley S.D.; Shea T.P.; Benito M.-I.; Town C.D.;
RA Fujii C.Y.; Mason T.M.; Bowman C.L.; Barnstead M.E.; Feldblum T.V.;
RA Buell C.R.; Ketchum K.A.; Lee J.-J.; Ronning C.M.; Koo H.; Moffat K.S.;
RA Cronin L.A.; Shen M.; VanAken S.E.; Umayam L.; Tallon L.J.; Gill J.E.;
RA Adams M.D.; Carrera A.J.; Creasy T.H.; Goodman H.M.; Somerville C.R.;
RA Copenhaver G.P.; Preuss D.; Nierman W.C.; White O.; Eisen J.A.;
RA Salzberg S.L.; Fraser C.M.; Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC004165; AAC16956.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 331 AA; 35961 MW; A0DE93084B8C8BC3 CRC64;

Query Match 20.0%; Score 139; DB 10; Length 331;
Best Local Similarity 36.2%; Pred. No. 1.6e-06;
Matches 42; Conservative 15; Mismatches 41; Indels 18; Gaps 5;
QY 1 MQSLQDPDPAARYNVLEALWRIIRTEGLWPRMGLNVTATGAGPAHALYFACYEKLKTL 60
DB 66 MQALRSCP-IKPIGIRQAFRSIIKTGDSALYRGWGLGAGPAHAYFSFVSVSKFL 124
QY 61 SDVIHPGG--NSHTANGAACVATLLHDAAMNPAE-----GNDSSYHSVGSCT 106
DB 125 S-----GGNPNNSAAHAISGVFATISSDAVTPMDVMVKQLQIGN--GTYKGVWDC 173

RESULT 14
014281 PRELIMINARY; PRT; 303 AA.
AC 014281;
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE PUTATIVE MITOCHONDRIAL CARRIER C8C9.12C.
GN SPAC8C9.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2002, 00:33:30 ; Search time 1936.36 Seconds
(without alignments)
11833.831 Million cell updates/sec

Title: US-09-870-113-1

Perfect score: 1095

Sequence: 1 atggagttggaggcgggg.....agtggaggctggcaagtga 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_vt:*

14: gb_vt:*

15: gb_vt:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pi:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vt:*

30: em_hgt_hum:*

31: em_hgt_inv:*

32: em_hgt_other:*

33: em_hgtg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

RESULT 1

AF327402

LOCUS

DEFINITION

AF327402

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AF327402 1448 bp mRNA linear PRI 02-MAY-2001

Homo sapiens putative mitochondrial solute carrier splice variant

mRNA, complete cds, alternatively spliced, nuclear gene for

mitochondrial product.

AF327402 GI:13926046

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1448)

Li, F.Y., Nikali, K., Grogan, J., Leibiger, I., Leibiger, B.,

Schweyen, R., Larsson, C. and Suomalainen, A.

Characterization of a novel human putative mitochondrial

transporter homologous to the yeast mitochondrial RNA splicing

proteins 3 and 4

FEBS Lett. 494 (1-2), 79-84 (2001)

21195335

11297739

2 (bases 1 to 1448)

ALIGNMENTS

Db 182 CCGATGCAGAGTCTACAGCCTGACCCAGCTGCCCGCTATCGCAATGTGTTGGAGGCCCT 241
Qy 354 ctgagagattataagaacgaggccctatgagcccatgagggcctgaaagctcaagc 413
Db 242 CTGGAGGATTTATAGAACGAGGCGCTATGAGGCGCCATGAGGGGCTCAACGTCACAGC 301
Qy 414 aacaggcagggcctgcccacgccccttatttgcctgctacgaaagttaaaaaagac 473
Db 302 AACAGCGCAGGGCTGCCACGCCCTTTATTTGCTGCTACGAAAGTTAAAAAGAC 361
Qy 474 attgaagtatgaataccaccctgggggcaatagccatatcccaatgctgcccgggtg 533
Db 362 ATTGAGTGATGTAATCCACCTCGGGGCAATAGCCATATGCCAATGTGCGCGGGGT 421
Qy 534 tctggaacattactcatgatgagccatgaacctcggaagtggcgaagcagagat 593
Db 422 TGTGACACATTACTTCATGATGACGCCATGAACCTCGGAAAGTGGTCAGCAGAGAT 481
Qy 594 gcagatgtacaactcaccatcaccacccgggtgacagactgtgtacgggcagtgtggcaaaa 653
Db 482 GCAGATGTACAACCTCACCATACACCGGTGACAGACTGTGTACGGCAGTGTGGCAAAA 541
Qy 654 tgaaggccggggccctttaccgcagctacacacccagctgaccatgaacgttccctt 713
Db 542 TGAAGGGCCGGGGCTTTTACCGCAGCTACACCCAGCTGACCATGAACGTTCCCTTT 601
Qy 714 ccaagccattcactcatgacctatgaattctcagagcaccttaacccccagagacg 773
Db 602 CCAAGCCATTCACTTCATGACCTATGAATTCCTGCAGAGGACACTTTAACCCCGAGAGC 661
Qy 774 gtacaacccaagctccacgtcctctctgtgagcttgccagagagctgtagctgcgcgac 833
Db 662 GTACAACCCAGCTCCACGCTCTCTGTGAGCTGTCGCCAGAGCTGTAGCTGCCGAGC 721
Qy 834 caaaccccaactgagcgtttgcaaaaactgctcaaacccagagagctcctggtttgaa 893
Db 722 CACAACCCACTGGGACGTTTGCAAAACACTGCTCAACACCCAGGAGTCTTGGCTTTGAA 781
Qy 894 ctcacacattacagacatatcacagcagctagcttagctcctcagagcaggtatatcaagt 953
Db 782 CTCACACATTACAGAGCATATCACAGGATGGCTAGTGCCTTCAGACGGTATATCAAGT 841
Qy 954 aggtgggtgacccgctatttccagggggtgagggccagagtaatttaccagatccccctc 1013
Db 842 AGGTGGGTGACCGCTATTTCCGAGGGGTGAGGCCAGAGTAATTTACAGATCCCTC 901
Qy 1014 cacagccatcgatgctgtgtatgattcttcaaaatacctaataacaaagcaaga 1073
Db 902 CACAGCCATCGCATGGTGTGTATGAGTTCTCAAAATACCTAATCACTAAAGGCAAGA 961
Qy 1074 agagtggagggtggcaagtga 1095
Db 962 AGAGTGGAGGGCTGCAAGTGA 983

RESULT 4
AF327403 AF327403 1889 bp mRNA linear PRI 02-MAY-2001
LOCUS Homo sapiens putative mitochondrial solute carrier splice variant
DEFINITION mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.
ACCESSION AF327403
VERSION AF327403.1 GI:13926049
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1889)
AUTHORS Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B.,
Schweyen, R., Larsson, C. and Suomalainen, A.
TITLE Characterization of a novel human putative mitochondrial
transporter homologous to the yeast mitochondrial RNA splicing

proteins 3 and 4
FEBS Lett. 494 (1-2), 79-84 (2001)
21195335
11297739
2 (bases 1 to 1889)
Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B., Schweyen, R.,
Larsson, C. and Suomalainen, A.
Direct Submission
Submitted (08-DEC-2000) Human Molecular Genetics, National Public
Health Institute, Mannerheimintie 166, Helsinki 00300, Finland
JOURNAL
FEATURES
source Location/Qualifiers
1..1889 /organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q24"
1032..1565
/note="similar to mitochondrial RNA splicing protein 3/4;
alternatively spliced"
/codon_start=1
/product="putative mitochondrial solute carrier splice
variant"
/protein_id="AAK49520.1"
/db_xref="GI:13926050"
/translation="MNPAEVVKQRMQNPSPHYRVTDCVRAVQNEGAGAFYSYTTQ
LTNVPFQAIFHMTYEFLEQHEFNPPSSHLVSGACGAVAAAATFDVCKTLK
NTQESLALNSHIITGHITGMASAFRTYVQGVGTAYFRGVQARVIYQIPSTAIAMSVYE
FFKYLITKQBEWRAGK"
BASE COUNT 404 a 508 c 544 g 433 t
ORIGIN
Query Match 56.7%; Score 621.4; DB 9; Length 1889;
Best Local Similarity 70.3%; Pred. No. 2.1e-101;
Matches 1094; Conservative 0; Mismatches 1; Indels 462; Gaps 1;
Qy 1 atgagattggaggggcggggtgctggcggtgtgcgggggggcgggcgggcgggcgggcg 60
Db 9 ATGGAGTTGGAGGGCGGGGTGCTGGCGGTGTGGCGGGGGGCGAGCGCGCGGGG 68
Qy 61 cggagccccggggagtcgagcgtgctgagcgggtgctgagcggggcggtgagcggggcg 120
Db 69 CGGAGCCCCGGGGAGTCGGCGCTGCTGGACGGGTGGCTGACGGGGGGCTGGCGGGGG 128
Qy 121 gccggcgggggaggccggggcctgagggccccgggttacgacaagatccggactccggc 180
Db 129 GCCGCGCGGGGAGCGCGGGCTGTCAGGCCCGCGGTACGACAAGATCCGACTCCGCGC 188
Qy 181 ccggactacggggcggtccggcgctggagccactgtcaccacgacacatggtggcagggcc 240
Db 189 CCGGACTACGAGGGCGCTCGCGGCTGGAGCCACTGTGCACACGACATCGTGCGACGCC 248
Qy 241 gtggcagggatcctggagcactgcgtgtgtagtaccaccatcgactgcgtcaagaccggatg 300
Db 249 GTGGCAGGGATCCTGGAGACACTGCGTGATGATACCCCATCGACTGCGTCAAGACCGGATG 308
Qy 301 cagagtctacagcctgacccagctgcccgcctatcgctcgaattgtgtggagggccctctggagg 360
Db 309 CAGAGTCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTTGGAGGCCCTCTGGAGG 368
Qy 361 attataaagacggaggcgctatgagggccatgagggggctgaaagctcacagcaaacgac 420
Db 369 ATTATAAGAACGGAGGCGCTATGGAGGCCCATGAGGGGGCTGAACGTCAACCAACAGCG 428
Qy 421 gcaggcgctgccacgccccttatttgcctgctacgaaagttaaaaaagacattgagt 480
Db 429 GCAGGGCTGCCCGCGCCCTTTATTTGCTGCTACGAAAGTTAAAAAGACATTGAGT 488
Qy 481 gatgtaatccaccctgggggcaatagccatatgtgccaat----- 519
Db 489 GATGTAATCCACCTGGGGCAATAGCCATATGGCAATGGTATTGAGCCTTCCTGTGCT 548
Qy 520 ----- 519


```
|||||
Db 309 CAGAGTCTACAGCCTGACCCAGCTGCCCGCTATGCAATGTTGGAGGCCCTCTGGAGG 368
Qy 361 attataaagcagggcctatgagggcccatgaggggctgaacgtcacagcaacaggc 420
Db 369 ATTATGAAGAGGGAGGGCTATGGAGGCCCATGAGGGGGCTGAACGTCAACACAGGC 428
Qy 421 gcaggcctgccagcgcctttatttctgctacgaaaaagttaaaaaagacattgagt 480
Db 429 GCAGGGCCTGCCACGCCCTTATTTGGCTGCTACGAAAGTTAAAAAGACATTGAGT 488
Qy 481 gatgtaatccacctgggggcaatagccatattgccaat----- 519
Db 489 GATGTAATCCACCCTGGGGCAATAGCCATATTGCAATGGTATTGAGCCTTCCTGTGCT 548
Qy 520 ----- 519
Db 549 GGTCCCCCACTTTCCCAACTCTTTGGGCTTTGTGCTGTGCTCAGTGCTTTCCAGTCTCAGC 608
Qy 520 ----- 519
Db 609 ATGGTTGGAGCTGAAGCTTTGGGCTGGGATAGGCCACTGGTGCGCAGTGATGCTGGC 668
Qy 520 ----- 519
Db 669 ATCCCTAGCAGCAGTATCTGCTTCTGCTGGGGTGTGAGCTGCATTTATTCTCAGA 728
Qy 520 ----- 519
Db 729 ATGATCTTTATTGATAGACTTGAAGCTGAGCTGGCCTTCCTATCATCYGGATGTGGAATACATTAGT 788
Qy 520 ----- 519
Db 789 GACCTTACAAAGTTGTTGGGAACAGATACTTTTACCTTCTTAAACAGAGTTTAGGAGCAG 848
Qy 520 ----- 519
Db 849 TGGGTCCCCATCTTTTGGACTAGCTCTTAACGTTACTTTTCCCGCTGTAGTGTAGCACA 908
Qy 520 ----- 519
Db 909 GCCACTCCCTTCACTGGGGGACCTCAGTGAGTTGGTCAGCTCTCTTGGCCCTTACATGTG 968
Qy 520 ----- 519
Db 969 GCAGTTGTTTTCTTGTGTCAGGTGCGGCGGGTGTGGCAACATTTACTTCATGATGCA 1028
Qy 559 gccatgaacctgcggaagtggtcaagcagagagatgcagatgtacaactcaccataccac 618
Db 1029 GCCATTGAACCCCTCGGAAGTGTCAAGCAGAGAGATGCAGATGTACAACCTCACCATACCAC 1088
Qy 619 cgggtgacagactgttacggcagtggtgcaaatgaaggcgccggccctttaccgc 678
Db 1089 CGGGTGACAGACTGTGTACGGCAGTGTGGCAAAATGAAGGGCGCGGGCCCTTTACCGC 1148
Qy 679 agtacacacacccagctgaccatgaactgctcttcttccaaagccattcatgacctat 738
Db 1149 AGCTACACACCCAGCTGACCTGACCTGAAAGCTTCTCTTCCAAAGCCATTACATTCAGCTAT 1208
Qy 739 gaattcctcagagacactttaacccagagacaggtacaaccaagctccacgctcc 798
Db 1209 GAATTCCTGCAGAGACATTTTAACCCCCAGACGGGTACAACCAAGCTTCCACAGCTCTC 1268
Qy 799 tctggagcttgcgagagactgtagctgcgcgagccacacacccactggacgttttgc 858
Db 1269 TCTGGAGCTTGGCAGAGACTGTAGTGCAGCCACAAACCCACTGACGCTTTGCAAA 1328
Qy 859 acaactgctcaacacccagagctcttggctttgaactcacacattacaggaacatatca 918
Db 1329 ACACCTGCTCAACACCCAGAGCTCTTGGCTTTGAACCTCACATTTACAGGACATATCACA 1388
Qy 919 ggcattgctagtccttcagagcggtatataagtagtgagggtaccgcctatttccga 978
|||||
```

```
Db 1389 GGCATGGCTAGTGCCTTCAGGACGGTATATCAAGTAGTGGGTGACCGCCTATTTCGGA 1448
Qy 979 ggggtgcaggccagagtaattaccagatccctccacagccatcgatgctgtgtat 1038
Db 1449 GGGGTGCAGGCGCAGAGTAATTTACCAGATCCCTCCACAGCCATCGCATGGTGTGTAT 1508
Qy 1039 gagttcttcaataacctaatcactaaaggaagaagtagtgagggtcgcaagtga 1095
Db 1509 GAGTTCTTCAAAATACCTAATCATTAAAAAGCAAGAGTAGTGGAGGCTGGCAAGTGA 1565

RESULT 6
AF267854 AF267854 1244 bp mRNA linear PRI 02-JAN-2001
LOCUS Homo sapiens NP016 mRNA, complete cds.
DEFINITION Homo sapiens NP016 mRNA, complete cds.
ACCESSION AF267854
VERSION AF267854.1 GI:12006034
KEYWORDS FLI_CDNA.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1244)
AUTHORS Xu X., Yang X., Gao G., Xiao H., Chen Z. and Han Z.
TITLE Direct Submission
JOURNAL Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, People's Republic of China
FEATURES
Source 1. 1244
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="normal pituitary"
381..911
/codon_start=1
/product="NP016"
/protein_id="AAG44723.1"
/db_xref="GI:12006035"
/translation="MALIVKORMQMNPSYHRVDCVRVWQNEGAGAFYRSYTTOL
TMNYPQAIHMTYEFLOEHFNPPORRYNPSHVSAGACAVAAAATPLDVCFTLLN
TQESIALNSHITGHITGMASAFRTYVQGGVTAYFRGVQARVIYIIPSTAIANSVYEF
FKYLITKROEEMRAGK"
BASE COUNT 308 a 331 c 290 g 315 t
ORIGIN
Query Match 47.5%; Score 520; DB 9; Length 1244;
Best Local Similarity 100.0%; Pred. No. 2.8e-83;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 576 agtgtcagcagagagatgcagatgtacaactcaccataccacccgggtgacagactgtgt 635
Db 392 AGTGTGTCAGCAGAGAGATGCAGATGTACAACCTACCACCCGGGTGACAGACTGTGT 451
Qy 636 acggcagtgtagcaaatgaaggcgccgttttaccgcagctacacacccagct 695
Db 452 ACGGCAGTGTGGCAAAATGAAGGGCGCGGGCTTTTACCGCAGCTACACCCAGCT 511
Qy 696 gaccatgaactgtcttttccaaagccattcacttcattgcactatgaattcctgcaggagca 755
Db 512 GACCATTGAACGTTCTCTTCCAAAGCCATTCTTTCATGACCTATGAATTCCTCGAGGACA 571
Qy 756 ctttaaacccagagacggtacaaccaagctccacgctctctctggagcttggcagag 815
Db 572 CTTTAAACCCAGAGAGCGGTACAACCAAGCTCCCAAGCTCTCTCTGAGAGCTTGGCAGG 631
Qy 816 agctgtagctgcgcgagccacacacccactggacgttttgcacaaactgctcaacacca 875
Db 632 AGCTGTAGTGTGGCAGCCACAAACCCACTGGAGGCTTTGCAAAACACTGCTCAACACCA 691
Qy 876 ggagtccttggctttgaactcacacattacaggaacatatcacaggcatggcttagtcctt 935
Db 692 GGAGTCTCTGGCTTTTGAACCTCACATTTACAGGACATATCACAGGATGGCTAGTGCCTT 751
```

Qy 936 caggacggtatataagtagtggtgacgcgcctattccgaggggtgcagggccagagt 995
 |||||
 Db 752 CAGGACGGTATATCAAGTAGTGCGGTGACCGCCTATTTCGAGGGGTGCAGGCCAGAGT 811
 |||||
 Qy 996 aattaccagatccctccacagccatcgcatgctgtgtatgatgattcttcaaatacct 1055
 |||||
 Db 812 AATTACAGATCCCTCCACAGCCATCGCATGCTGTGTATGAGTTCTTCAATACCT 871
 |||||
 Qy 1056 aatcactaaagcaagcaagtaggagggctggaagtga 1095
 |||||
 Db 872 AATCACTAAAAGGCAAGAGTGGAGGGCTGGCAAGTGA 911
 |||||

RESULT 7
 AL353719/c
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete sequence.
 ACCESSION AL353719 AC007643
 VERSION
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 123160)
 Ramsay,H.
 Direct Submission
 Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonequery@sanger.ac.uk

On Sep 26, 2001 this sequence version replaced gi:14280413.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em:, EMBL; Sw:,
 SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information on the WORMPEP
 database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr10
 RP11-85A1 is from the library RPCI-11.1 constructed by the group of
 Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBac3.6

IMPORTANT: This sequence is not the entire insert of clone
 RP11-85A1 it may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true left end of clone RP11-85A1 is at 1 in this sequence. The
 true left end of clone RP11-483F11 is at 123061 in this sequence.
 The true right end of clone RP11-129J12 is at 51589 in this
 sequence.

FEATURES

Location/Qualifiers
 1. .123160
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-85A1"
 /clone_lib="RPCI-11.1"

misc_feature complement(7066..7118)
 /note="Single clone region. Sequence from reads from a
 short insert library derived from a single pUC clone.
 Restriction digest data confirm the assembly."
 52162
 unsure
 BASE COUNT 33221 a 28084 c 27196 g 34659 t
 ORIGIN

Query Match 47.5%; Score 520; DB 9; Length 123160;
 Best Local Similarity 100.0%; Pred. No. 1.6e-83;
 Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 576 agtgcacagcagagatgacatgacactcaccatcacccgggtgcagactgtgt 635
 |||||
 Db 82990 AGTGGTCAAGCAGAGAGATGCAGATGTACACTACCATACCCGGGTGCAGACTGTGT 82931
 |||||
 Qy 636 acggggcagtggtggcaaaatgaagggggcggccttttaccgcagctacacacccagct 695
 |||||
 Db 82930 ACGGCGAGTGTGGCAAAATGAAGGGCGCGGCGCTTTTACCGAGCTACACCCAGCT 82871
 |||||
 Qy 596 gaccatgaacgtctctttccaaagcattcaactcatgacctatgaattcctgcaggagca 755
 |||||
 Db 82870 GACCATGAACGTTCTTTTCCAAAGCCATTACCTTTCATGACCTATGAATTCCTGCAGGAGCA 82811
 |||||
 Qy 756 ctttaacccccagagcaggtacaaacccaagctccacagctcctctgtgagcttgcgcaag 815
 |||||
 Db 82810 CTTTAAACCCCGAGAGCGGTACAAACCCAGCTCCACGCTCTCTCTGGAGCTTGGCGAGG 82751
 |||||
 Qy 816 agctgtagctgccgcagccacacacccactggacgtttgcaaaacactgctcaacaccca 875
 |||||
 Db 82750 AGCTGTAGTGCCTGCAGCAGCACACCCACTGCGAGTTTGCAAAACACTGCTCAACACCCA 82691
 |||||
 Qy 876 ggaagctctgtgcttggaactcacacattacagagacatatacacagcagctgtagtgcctt 935
 |||||
 Db 82690 GGAGTCTCTGGCTTTGAACTCACACATTACAGGACATATACAGGSCATATCAGGSCATGCGTCT 82631
 |||||
 Qy 936 caggacggtatataagtagtggtgacgcgcctattccgaggggtgcagggccagagt 995
 |||||
 Db 82630 CAGGACGGTATATCAAGTAGTGCGGTGACCGCCTATTTCGAGGGGTGCAGGCCAGAGT 82571
 |||||
 Qy 996 aattaccagatccctccacagccatcgcatgctgtgtatgatgattcttcaaatacct 1055
 |||||
 Db 82570 AATTACAGATCCCTCCACAGCCATCGCATGCTGTGTATGAGTTCTTCAATACCT 82511
 |||||
 Qy 1056 aatcactaaagcaagcaagtaggagggctggaagtga 1095
 |||||
 Db 82510 AATCACTAAAAGGCAAGAGTGGAGGGCTGGCAAGTGA 82471
 |||||

RESULT 8
 AK056782
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ32220 fis, clone PLACE6003850, moderately similar to Homo sapiens mitochondrial solute carrier mRNA.
 ACCESSION AK056782
 VERSION
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens placenta cDNA to mRNA, clone_lib:PLACE6 clone:PLACE6003850.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (sites)
 Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
 Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
 Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
 Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
 Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,
 Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
 Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,

Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1588)
 Direct Submission
 Isogai, T., Otsuki, T. and Sugiyama, T.
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, tel: 81-438-52-3951, fax: 81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.

FEATURES
 source
 1. .1588
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLACE6003850"
 /tissue_type="placenta"
 /clone_lib="PLACE6"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 385 a 405 c 375 g 423 t
 ORIGIN

Query Match 47.3%; Score 518.4; DB 9; Length 1588;
 Best Local Similarity 99.8%; Pred. NO. 5.2e-83;
 Matches 519; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 576 agtgcacagagatgcagatgatacactcaccatcacccgggagacagtgtgt 635
 |||
 Db 745 AGTGGTCACGAGAGATGCAGATGTACACTCACCATTACCCGGGGACAGGCTGTGT 804
 QY 636 acgggcagtggtgcacaaatgaaggggcgccgttttaccgcagctacacacccagct 695
 |||
 Db 805 ACGGGCAGGTGTGGCAAAATGAAGGGCGCGGGCCCTTTACCCAGCTACACCCAGCT 864
 QY 696 gaccatgaacgtctcttcacagcattcaactcattgatgacctgaattcctcagagaca 755
 |||
 Db 865 GACCATGAACGTTCTCTTCCAGGCCATTACATTCATGACCTATGAATTCCTCGCAGGACA 924
 QY 756 cttaacccccagagcaggtatacaccacgaagctccacagctcctctctggagcttgcgcag 815
 |||
 Db 925 CTTTAACCCAGAGACGGTACACCCAGCTCCACGCTCTCTCGGAGCTTGGCAGG 984
 QY 816 agctgtagctgcgcagccacacccccactggagcgttttgcaaaacactgctcaaacacca 875
 |||
 Db 985 AGCTGTAGTGGCGGACGACACACCCCACTGTGACGCTTTGCAAAACACTGCTCAACACCCA 1044
 QY 876 gtagctctgcttgaactcacattacagacacatcacagcagcatgctgagcttgcctt 935
 |||
 Db 1045 GGAGTCCTTGGCTTTGAACATCACATACAGGACATATACAGGATGCGCTAGTGCCTT 1104
 QY 936 caggacggtatatacagtagtggtgggtgaccgcctatttccggaggggtgcagccagagt 995
 |||
 Db 1105 CAGGAGGTTATATCAAGTAGTGGGTGACCGCCCTATTTCGAGGGGTGCAGCCAGAGT 1164
 QY 996 aattaccagatccctccacagccatgcagctggtctgtgtatgagtttccaaatacct 1055
 |||
 Db 1165 AATTTACCAGATCCCTCCACAGCCATCGATGGTCTGTGTATGAGTTCTTCAATAACCT 1224
 QY 1056 aatcactaaagcgaagagagtgaaggctgcgaagtga 1095
 |||
 Db 1225 AATCACTAAAGGCAAGAGAGTGGAGGCTGGCAAGTGA 1264

RESULT 9
 BC023172
 LOCUS
 1072 bp mRNA linear ROD 07-FEB-2002

DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT

Mus musculus, Similar to putative mitochondrial solute carrier,
 clone MGC:37028 IMAGE:4949779, mRNA, complete cds.
 BC023172
 BC023172.1 GI:18606247
 MGC.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1072)
 Strausberg, R.
 Direct Submission
 Submitted (04-FEB-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAC Plate: 60 Row: h Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES
 source

1. .1072
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:37028 IMAGE:4949779"
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma 5 month old virgin mouse."
 /clone_lib="NCI-CGAP_Mam6"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 188. .721
 /codon_start=1
 /product="Similar to putative mitochondrial solute
 carrier"
 /protein_id="AAH23172.1"
 /db_xref="gi:18606248"
 /translation="MNPAAVYVKRMQWNSYHRVTCVRAVWQNEGAGAFYRSYTTQ
 LTMVFFOAIHEFTYFLOEFHFNPNRYPNSHVLGACAGADAATPLDVKTKLL
 NTOESLANSNITGHITGMASAFRTVYQGVGVAFVGVQARVIYQIPSTAIAWSVE
 FFKYLITKQEEWRAGK"
 BASE COUNT 268 a 308 c 263 g 233 t
 ORIGIN

Query Match 46.0%; Score 504; DB 10; Length 1072;
 Best Local Similarity 92.2%; Pred. No. 2e-80; Mismatches 45; Indels 0; Gaps 0;
 Matches 531; Conservative 0;

QY 520 ggtgcggcggtgtgtggaacattacttcattgatgcagccatgaacctgcggaagt 579
 |||
 Db 146 GGTGACGCGGATGTGTGCGGACATTTACTTCATGATGAGCCATGAATCCACGCGAAGTC 205
 QY 580 gtcaagcagagatgcagatgtacaactcaccatccaccgggtgacagactgtgtaagg 639
 |||
 Db 206 GTCAAGCAGAGATGCAGATGTACAACCTGCCGTACCCCGGTGACAGACTGTGTTCGG 265

Db 99150 AGCTGTAGTCCCGCCGCCACCCACTGACGCTTGTGCAAAACACTGCTCAACACCCA 99091
Qy 876 ggaagtccttggttgaactcacacattacagacatatacacaggcatggttagtgcctt 935
Db 99090 GGAGTCCCTGGCCTTGTAGCTCAAAACATTACAGACACATACAGGCATGGCTAATGCTT 99031
Qy 936 caggacggatatacaagtagtggtgagccgcatttccaggggtgcagggccagag 995
Db 99030 CAGGACGGTCTATCAAGTAGGGGGGTGACTGCTTACTTCCAGGGGTACAGGCAGAGT 98971
Qy 996 aattaccagatccctccacagccatcgcatggtctgtgtatgagttcttcaataacct 1055
Db 98970 AATTACCAGATCCCTCCACAGCCATCGCATGGTCTGTGTATGAATCTTCAAATACT 98911
Qy 1056 aatcactaaaggcaagagtggtgaggggtgcgaagtgga 1095
Db 98910 AATCAGGAAGCGGCAAGAAGTAATGAGGGCAGGCAAGTGA 98871
RESULT 11
AF288621 4112 bp mRNA linear ROD 27-OCT-2001
LOCUS Mus musculus mitochondrial solute carrier-like protein mRNA,
DEFINITION complete cds.
ACCESSION AF288621
VERSION AF288621.1 GI:16506177
KEYWORDS SOURCE
ORGANISM house mouse.
REFERENCE 1 (bases 1 to 4112)
AUTHORS Li,Q.-Z., Ruan,Q.-G., Eckenrode,S., Shi,J.-D., Cruz,P., Wang,C.-Y.
and She,J.-X.
TITLE A new gene which is highly expressed in NOD mice spleen
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 4112)
REFERENCE Li,Q.-Z., Shi,J.-D., Cruz,P., Echenrode,S., Ruan,Q.-G. and
Wang,C.-Y.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2000) Pathology, University of Florida, 1600 SW
Archer Road, Rm. D6-15, Gainesville, FL 32610, USA
FEATURES
Location/Qualifiers
1..4112
/organism="Mus musculus"
/db_xref="taxon:10090"
116..1132
/note="50E12"
/codon_start=1
/product="mitochondrial solute carrier-like protein"
/protein_id="AAI23859.1"
/db_xref="GI:16506178"
/translation="MELRRGGVQAGRRMDGCDRGCGSKDAGSDYENLPTFAS
VSTHMTACAMAGILEHSTYVDSVKRMQSLNPDPKARYTSYIGALKRIMHTEGFWR
PLRLNVMWAGPAHAMFYACENMKRTLNDVFSHOGNSHLANGVAGSMATLLHDV
MNPBEVKQRLQMYNSHQSAFSCRTVWTEGFAFYRSYVLTQMTNIPQSLIHFT
YEFQEQVNPREDINPQSHIISGGLAGALAAAATPLDYCKTLNTQENMALSLANVS
GRLSGMANAFRTVYQLNGLAGYFKGIQARVIYQMPSTAISWSVYEFFKYLTKROLN
RTLY"
BASE COUNT 970 a 1024 c 970 g 1148 t
ORIGIN
Query Match 41.6%; Score 455.8; DB 10; Length 4112;
Best Local Similarity 67.1%; Pred. No. 6.9e-72;
Matches 678; Conservative 0; Mismatches 327; Indels 6; Gaps 2;
Qy 73 gagtcgagcgtctgagcagggtggctgagcggggtggcggtggcgggggggcgcgcgcg 132
Db 104 GCGGGGCCCGGATGAGTCCCGCTGGCGGTGGGGAACACAGCTCGGGCGCGGAGG 163
Qy 133 gaggccgggggctcgaggcccccggtacgacaagatccgagcctcgcccc--ggactac 189

* 152125 152224: gap of unknown length
* 152225 158778: contig of 6654 bp in length
* 158779 158978: gap of unknown length
* 158979 163928: contig of 4950 bp in length
* 163929 164028: gap of unknown length
* 164029 169353: contig of 5325 bp in length
* 169354 169453: gap of unknown length
* 169454 175590: contig of 6137 bp in length
* 175591 175690: gap of unknown length
* 175691 180475: contig of 4785 bp in length
* 180476 180575: gap of unknown length
* 180576 184657: contig of 4082 bp in length
* 184658 184757: gap of unknown length
* 184758 190332: contig of 5575 bp in length
* 190333 190432: gap of unknown length
* 190433 194134: contig of 3702 bp in length
* 194135 194234: gap of unknown length
* 194235 196592: contig of 2358 bp in length
* 196593 196692: gap of unknown length
* 196693 200530: contig of 3838 bp in length
* 200531 202809: contig of 2179 bp in length
* 202810 202909: gap of unknown length
* 202910 205034: contig of 2125 bp in length
* 205035 205134: gap of unknown length
* 205135 207176: contig of 2042 bp in length
* 207177 207276: gap of unknown length
* 207277 208845: contig of 1569 bp in length
* 208846 211271: contig of 2326 bp in length
* 208946 211271: gap of unknown length
* 211272 213302: contig of 1931 bp in length
* 213303 213402: gap of unknown length
* 213403 215484: contig of 2082 bp in length
* 215485 215584: gap of unknown length
* 215585 216969: contig of 1385 bp in length
* 216970 217069: gap of unknown length
* 217070 218358: contig of 1289 bp in length
* 218359 218458: gap of unknown length
* 218459 219886: contig of 1428 bp in length
* 219887 219986: gap of unknown length
* 219987 221062: contig of 1076 bp in length.
FEATURES
Location/Qualifiers
1..221062
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/chromosome="Rf1"
/clone="CH230-24M6"
BASE COUNT 58389 a 50577 c 49732 g 58877 t 3487 others
ORIGIN
Query Match 42.7%; Score 457.2; DB 2; Length 221062;
Best Local Similarity 93.7%; Pred. No. 3.8e-74;
Matches 487; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Qy 576 agtggcagcagagagatgcagatgtacacatcacaccacccgggtgacagactgtgt 635
Db 99390 AGTGGTCAAGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 99331
Qy 636 acggcgagctgtggcaaaatgaaggggccggggcctttaccgcagctacaccaccagct 695
Db 99330 TCGGGCAGTGTGCAAAATGAAGGGGCCGGGCTTTTACCAGCTACACACCCAGCT 99271
Qy 696 gaccatgaagcttcttcccaagcattcactcatgactatgaattcttcagcagagca 755
Db 99270 GACCATGAATGTTCCTTCCAAAGCCATTTCATTCATGACTATGATGATGATGATGAT 99211
Qy 756 ctttaacccccagagagcgttacaacccaagctcccaagctctctctctgtgagcttgagcag 815
Db 99210 CTTTAATCCCAAGAGCGGTACAAACCCAGCTCCCAATGTGCTGTGTGAGCCTGCGCAGG 99151
Qy 816 agctgtagctgcgcagcgcacacacccactgagcgttttgcacaaactgctcaacaccca 875

```
Db 164 ATGATGGGACTGCGCGATGCGGCTGCGGCAGCAAGGACGCGGTCGAGGACTAC 223
Qy 190 gagcgctgcggtgagccactgtcaccaacacacatggtggcagcgccgtggcagg 249
Db 224 GAGAACCTACCGACGCGCTCGGTTCCTCCACACATGACAGCGGGAGCGATGCCGG 283
Qy 250 atcttgagcactcgtgatgtacccatcgactcgtcgaagaccgagatcagagctca 309
Db 284 ATTCTGAGCACTCCATCATGATACCGGTGACATCCGTGAACACAGGATGACAGTTG 343
Qy 310 cagcctgaccacgctgcccctcgaatgtgttgaggccctctgagagctataaga 369
Db 344 AATCCAGATCCAAAGCCCGGTATACAGCATCTATGCGGCCCTCAAGAGGATCATGCAC 403
Qy 370 acgagggccctatgagggcccatgaggggctgaacgtcacagacaacagcgagggcct 429
Db 404 ACTAAGGCTTCTGGAGGCCCTCGCGGGCTGAACCTGATGATGGGTGACAGGCC 463
Qy 430 gccacgcccctttatttgcctgtacgaaagttaaataagacattgagtgttaact 489
Db 464 GCGACGCCATGATTTGCTGCTATGAACACATGAAGGACTTTAATGACGTTTTC 523
Qy 490 caccctggggcaatagccatttcccaatggtgcggccgggtgtgtggaacattactt 549
Db 524 AGCCACCAAGGAACAGCCATCTAGCTAATGGGTAGCTGGAGTATGGCCACCCCTACTC 583
Qy 550 catgatcagccatgaacctgcgaagtgtcaagcagagagatgcagatgataactca 609
Db 584 CACGATCAGTAATGATTCAGCAGAGAGTGGTGAACAGCGCTTACAGATGTACAACTCC 643
Qy 610 ccataccacgggtgcagacgtgtgacggcgagtggcaaatgaaggccggggcc 669
Db 644 CAGCACGATCAGCCTTCAGTTGTTATCCGGACAGTGTGGCGGACCGAGGGTTGGGGCC 703
Qy 670 tttaacgcagctacacccacccagctgaccatgaaccttcccttcccaagccattcctc 729
Db 704 TTCTACAGGAGTTACACACACAGCTGACCATGAATATCCCTTCCAGTCAATTCACCTC 763
Qy 730 atgacctataattctgcaggagcactttaacccccagagaggtacaacccaagctcc 789
Db 764 ATCACCATTAGTTTCTGAGGAGCAAGTCAACCCCTCGCGGGACTACACCCACAGCTCT 823
Qy 790 cagctctctctgaggtctgcaggagctgtagctgcgcagccacacaccccaactggac 849
Db 824 CACATCATCTCAGGAGCCCTGGCGGGGCACTGSCCGCAGCTGCCACCCCGCTGGAC 883
Qy 850 gtttgaaaaactgctcaacacccagaggtccttgggttgaaact---cacacattaca 906
Db 884 GTCTGCAAAACCCCTCCCTCAACACGACGAGAACATGCTCTCTCCCTGGCCACGTCAGC 943
Qy 907 ggacatacacagcaggtgtagtgcttcaggacggtatatcaagttagtggtgggtgacc 966
Db 944 GCGCGGTCTCGGGCATGGCCATGCCCTTCGGACGGTGTACAGCTCAACGGCTTGCC 1003
Qy 967 gcctattccgaggggtgcaggccagagtagtaattaccagatccccctccacagccatgca 1026
Db 1004 GGCTATTTCAAGGCATCCAGGCTCGAGTCACTTACCAGATGCCCTCCACCGCATCTCC 1063
Qy 1027 tggctgtgtatgagttcttcaataactaactaactaaaaggcaagaag 1077
Db 1064 TGGTCCGCTTTATGAGTCTTCAAGTACATCTTACAAAGAGGCGAGCTGGAG 1114
```

RESULT 12

```
AF223466 AF223466 1429 bp mRNA linear PRI 17-APR-2000
LOCUS Homo sapiens HT015 protein (HT015) mRNA, complete cds.
DEFINITION
ACCESSION AF223466
VERSION AF223466.1 GI:7578782
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1429)
Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
A novel gene expressed in human hypothalamus
Unpublished
REFERENCE
2 (bases 1 to 1429)
Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
Direct Submission
Submitted (12-JAN-2000) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, China
FEATURES
Location/Qualifiers
source
1..1429
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
1..1429
/gene="HT015"
45..1088
/gene="HT015"
/codon_start=1
/product="HT015 protein"
/protein_id="AAF64141.1"
/db_xref="GI:7578783"
/translaton="MELRSQVGSQAVRRMDGSDRGGGKDATGSEYENLPTSAS
VSHMTAGAMAGILESHVMPVDSVKRMOSLPDKAOTSYIGALKKIMRTEGWR
PLRGVNVIMINGAGPAHAMYFACYENMKRLINDVPHQGNHSLANGSLANGSLATLLHDV
MNPVEVVKQRLQMYNSOHSRISCIRTVWRTEGAGFYRITITOLTMNIPFQSIHFIT
YAEVQVNPHTYNPOSHISGLGALAAATPDLCKTLNLEQNVALSANIS
GRLSGMANAFRTYQLNGLPATSKASRRVSSRCPPPPFLGLSSSSSTFSPSASWKI
ELHTKGRDHRIFS"
BASE COUNT 326 a 426 c 402 g 275 t
ORIGIN

Query Match 39.9%; Score 436.4; DB 9; Length 1429;
Best Local Similarity 66.4%; Pred. No. 2.3e-68;
Matches 673; Conservative 0; Mismatches 336; Indels 5; Gaps 3;

Qy 61 cggagcccccggggagtcgagcgctgctggacgggtggtgctgcagcgggcgctggcggggg 120
Db 23 CTGCGCCCGCGGAGCTGGCGGATGAGCTGCGCAGCGGGAGCGTGGCGCAGCCAGCGCGGT 82
Qy 121 gccggcgcggggagcgcgggcctgcagcccccggtacgacaagatccggactccggc 180
Db 83 GGC CGC GAG GAT GAT GAT GGG GAC CAC CCG CAG AT G G C G C G G C G G C A A G G A C C C C G G G T C 142
Qy 181 -ccggactacgagcgctgcgcgctggagccactgtcaccacgcacatggtggcagcgcc 239
Db 143 GGAGGACTACGAGAACCTGCGGACTAGCGCTCGGTGTCACACCCACATGACAGCAGGAGC 202
Qy 240 cgtgcaggaatcctgagcactcgtgatgtaccctcactgactgcactgcgtcgaagaccggat 299
Db 203 GATGCGCGGATCCTGAGCAGCTCGGTCACTACCCGGTGGACTCGGTGAAGACACCAAT 262
Qy 300 gcagagtctacagctgacccagctgcgcctcgcgaatgtgttgaggccctctctggag 359
Db 263 GCAGAGTTTGAGTCCAGATCCCAAAGCCAGTACACAAATATCTACGAGGCCCTCAAGAA 322
Qy 360 gattataagacggagggccctatggagcccatagggggctgaggggctgacgtcacagcaacag 419
Db 323 AATCATCGGACCGAAGGCTTCTGGAGGCCCTTGGAGGCCCTTCAACGCTCATGATCATGGG 382
Qy 420 cgcagggcctgccacgcctttatttgcctgctacgaaagttaaataagacattgag 479
Db 383 TGCAGGGCCGCGCATCCATGATGTTTGGCTGCTATGATAAACATGATAAGGACTTTAAA 442
Qy 480 tgaatgaatcacccctggggccaatagccatattgccaatggtgcggcggtgtgtggc 539
Db 443 TGACGTTTTCCACCACCAAGCAACACACACCTAGCCACGGGATAGCTGGGATATGGC 502
Qy 540 aacattactcatgatgcagccatgaaccctgcggaagtgtgtaagcagaggtatgcagat 599
```

```

Db 503 CACCTGCTCCAGATCGCGTAATGATCCAGAGAGAGTGTGAAGCAGCGCTTGCGAGT 562
Qy 600 gtacaaactcaccataaccacccgggtgacagactgtgtacggcgagtggtggcaaaatgaag 659
Db 563 GTACAACTCGCAGCAGCGGTGAGCAATCAGCTGCATCGGACGGTGTGTGGAGACCGAGG 622
Qy 660 ggcggggccctttacgcagctacacacccagctgacacatgaacgttcccttccaaagc 719
Db 623 GTTGGGGCCCTTACCGGAGCTACACCGAGCTACACCGAGCTGACCATGAACATCCCGCTTCCAGTC 682
Qy 720 catcaactaatgacctatgaattccctgacagagcaactttaacccacagagacggtacaa 779
Db 683 CATCCACTTTCATCACTATGAGTTCTCTGAGGAGAGGTCAACCCCGACCGGACTACAA 742
Qy 780 cccaagctcccaagctccctctctgagcttgccgagagagctgtagctgcgcgagccacac 839
Db 743 CCGCAGTCCACATCATCTCAGCGGGGCTGCGCGGGCCCTGCCCGCGCGCCACGAC 802
Qy 840 cccactgacgttttgcaaaacactgctcaacacccagagtgctcttggttt---gaactc 896
Db 803 CCCCTGGACGCTCTGAAGACCTTCTGAACACTCAGGAGAACGTGGGCCCTCTCGCTGGC 862
Qy 897 acacattacagagacatatcaagcagcatggttagtgccttcaggagcggtatatcaagtagg 956
Db 863 CAACATCAGCGCGCGCTGTCCGGGTATGGCCAATGCCCTTCCGGAGGGTGTACCACTCAA 922
Qy 957 tggggtgacgcctattcccgaggggtgagcagccagagtaattaccagatccctccac 1016
Db 923 CGGCTG-CCGCTACTTCAAGAGCATCAGCGGGGTCTCATCTACCAAGATGCCCTCCAC 981
Qy 1017 agccatcgatggtgtgtatgagttcttcaaatcacctaatcaactaaagcca 1070
Db 982 CGCCATTTCTTGCTGTCTATGAGTTCTTCAAGTACTTCTTCAACCAAGCGCCA 1035

RESULT 13
AX306258
LOCUS AX306258 483 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 1009 from Patent WO0188188.
ACCESSION AX306258
VERSION AX306258.1 GI:17645513
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 1009 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
source
1. 483
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 127 a 139 c 121 g 96 t
ORIGIN

Query Match 38.2%; Score 417.8; DB 6; Length 483;
Best Local Similarity 93.2%; Pred. No. 5.5e-65;
Matches 437; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 462 gtaaaaaagacatgagtgatgtaattccacccctgggggcaatagccatattgccaatgg 521
Db 15 GTTAAAAAGACATTTGAGTACGTAAATCCACCCAGGGGCATATGCCATATTGCCAATGG 74
Qy 522 tgcggccgggtgtgtggcaacattacttcatgatgcagccatgaacccctgcggaagtggt 581
Db 75 TGCAGCCGGATGTGTGGCGACATTACTTCATGATGCAGCATGAATCCAGCGGAAGTCGT 134
Qy 582 caagcagagatgcagatgtatacactcaccataccacccgggtgacagactgtgtacgggc 641

```

```

Db 135 CAACGAGAGATGCAGATGTACAACTCGCCGTACACCGCTGACAGACTGTGTTCGGGC 194
Qy 642 agtgtgcaaaatgaagggccggggccctttttaccgagctacacccacccagctgacct 701
Db 195 AGTGTGGCAAAATGAAGGGCGCGGGCCCTTTTACCGAGCTTACAGACCCAGCTGACCAT 254
Qy 702 gaacgttcttccaaagccattcacttcacatgacctatgaattcctgcagagacactttaa 761
Db 255 GAATGTCCCTTCCAGCCATTCATCTCATGCCATGATGAGTCTCGAAGAGCATTATA 314
Qy 762 ccccgagagcgggtacacacccaaagctccacgctcctctctgagcttgcgcagagagctgt 821
Db 315 CCCCAGAGACGGTACACCCAGCTCCCATGTCTCTGTGGGCCCTCGCAGGAGCAGT 374
Qy 822 agctgcgcgagcccaacaccccaactggagcttgcacaaacactgctcaacacccaggaagtc 881
Db 375 AGTGTCCCGCCCAACACCCCACTGGAGCTTTGTAAAAACACTGTCTCAACACCCAGGAATC 434
Qy 882 cttggcttgaaactcaacacattacagagacatatcacagcagctagct 930
Db 435 CTGGCTTTGAACCTAAACATTACTGGACATCATCAGGCATGGCTAGT 483

RESULT 14
AX071440
LOCUS AX071440 401 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 1912 from Patent WO0102568.
ACCESSION AX071440
VERSION AX071440.1 GI:12581791
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 401)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J., Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D., Lamson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M., Labat,I., Leshkovitiz,D., Kita,D., Garcia,V. and Strache-Crain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 1912 11-JAN-2001;
CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
FEATURES
source
1. 401
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature
1. 401
/note="n = A,T,C or G"
BASE COUNT 100 a 104 c 120 g 76 t 1 others
ORIGIN

Query Match 35.8%; Score 392; DB 6; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.3e-60;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 gtctacagctgacccagctgcccctatcgcaatgtgttgaggccctctggaggatta 364
Db 9 GTCTACAGCTGACCCAGCTGCCCTATCGCAATGTGTGGAGGCCCTCTGGAGGATTA 68
Qy 365 taagaacgagggcccttgaggcccatgaggggctgaacgtcaacagcagcgccag 424
Db 69 TAAGAACGAGGGCCCTATGGAGGCCCATGAGGGGCTGAACGCTCAGACACACAGCGCAG 128
Qy 425 ggcctgccacccctcttatttgcctgctacgaaaagttaaaaaagacattgagtgc 484
Db 129 GGCCTGCCACGCCCTTTATTTTGTCTGCTACGAAAAGTTAAAAAAGACATTGAGTGATG 188
Qy 485 taatcacccctgggggcaatagccatattgccaatgtgtcgccggcggtgtgtggcaaat 544
Db 189 TAATCCACCCCTGGGGCAATAGCCATATTGCCAATGGTGGCGCGGGTGTGTGGCAACAT 248

```

```

Qy 545 tacttcgatgacgccatgaacctgcggaagtgttcaagcagagagtgatgagatgtaca 604
|||||
Db 249 TACTTCATGATGACCCATGACCTCGGAGTGGTCAACGAGAGGATGAGATGTACA 308
605 actcaccataccacccgggtgacagactgttacggcgagtggtgcaaaatgaaggggccg 664
|||||
Db 309 ACTCACCATACACCGGGTGACAGACTGTGACGGCGAGTGTGGCAAAATGAAGGGCGG 368
665 gggccttttacgcagactacacccacagctg 696
|||||
Db 369 GGGCCTTTTACCGCAGCTACACACCCACGCTG 400

RESULT 15
AF361699 5869 bp mRNA linear ROD 07-NOV-2001
LOCUS Mus musculus mitochondrial carrier-like protein mRNA, complete cds;
DEFINITION nuclear gene for mitochondrial product.
ACCESSION AF361699
VERSION AF361699.1 GI:16755527
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 5869)
AUTHORS Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and
She,J.
TITLE A novel mouse mitochondrial carrier protein gene is up-regulated
from young to adult NOD mice
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5869)
AUTHORS Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and
She,J.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2001) Department of Pathology, Immunology and
Laboratory Sciences, University of Florida, 1600 SW Archer Road,
Room D6-15, Gainesville, FL 32610, USA
FEATURES
source
1. 5869
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="14"
/map="between D14Mit124 and D14Mit236"
116..664
/note="MCLP"
/codon_start=1
/evidence=not_experimental
/product="mitochondrial carrier-like protein"
/protein_id="AAL27990.1"
/db_xref="GI:16755528"
/translation="MELRRGGVGNQAGRMDGDCRDGGCGSKDAGSEYENLPISAS
VSHMTAGAWAGLEHSIMIPVDSVTRMOSLAPDPDPKRYTSIYGALKRIMHTEGFWR
PLRLGNVMMGAPAHAMFYACYNMKRTLNDVFSHOGNSHLANGILKAFVWSWEALL
SGASSFGPSNLHPRTENSRTV"
BASE COUNT 1387 a 1427 c 1390 g 1665 t
ORIGIN

Query Match 25.7%; Score 281; DB 10; Length 5869;
Best Local Similarity 70.1%; Pred. No. 1e-40;
Matches 393; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy 520 ggtcgccgggtgtgtgcaaatcttctcatgatgcagccatgaacccctgcggaagtg 579
|||||
Db 2311 GGGTAGCTGGGATGTGGCCACCCTACTCCAGCATGCAGTAATGAATCCAGCAGAAGTG 2370
580 gtcaagcagagatgcagatgtacaactcaccatccacgggtgacagactgtgtacgg 639
|||||
Db 2371 GTGAACAGCGCTTACAGATGTACAACCTCCAGCACCAGTCAGCTTCAGTGTCCGG 2430
940 gcagtgtggcaaatgaaggggcggggccccttttaccgcagctacacccacccagctgacc 699

```

```

Db 2431 ACAGTGTGGCGGACGAGGGGTTGGGCGCTTCTACAGAGGTTACACACACAGCTGACC 2490
Qy 700 atgaacgttcctttccaagccattcaacttcattgaactgaattctctgcaggagcacttt 759
|||||
Db 2491 ATGAATATCCCGCTTCCAGTCAATTCACTTCACTATGAGTTTCTGCAGGAGCAAGTC 2550
760 aacccccagagagcgtacaacccaagctcccaacgtcctctctgagcttgcgcagagct 819
|||||
Db 2551 AACCTCGCGGGACTACAACCCACAGTCTCACATCATCTCAGGAGGCTGCGCCGGGGA 2610
820 gtactgcgcgagccacacccccactggacgtttgcaaaacactgctcaacacccaggag 879
|||||
Db 2611 CTGGCGCGAGCTGCCACCAACCCCGCTGGAGCTCTCAAAAACCTCTCTCAACACGAGGAG 2670
880 tctctggctttgaac---tcacacattacaggagacatatcacaggcatggctagtccttc 936
Db 2671 AACATGGGCTCTCTCCCTGGCCAACTCAGCGCGCGCTGTCGGGATGGCCAAATGCCCTTC 2730
937 aggcaggtatatcaagttaggtgggtgacgcctatttccgaggggtgcaggccagagta 996
Db 2731 CGGACGGTGTACCAGCTCAACGGCCTTGCCGGCTATTTCAAAGGCATCCAGGCTCGAGTC 2790
997 atttaccagatccctccacagccatcgatcggtgtgtgtatgatttcttcaaatcacta 1056
Db 2791 ATTTACCAGATGCCCTTCCACCGCCATCTCCTGGTCCGTTTATGAGTTCTTCAAGTACATC 2850
1057 atcactaaaaggcaagaagag 1077
Db 2851 CTTCAAAAGAGCGCAGCTGGAG 2871

```

Search completed: August 27, 2002, 02:34:02
Job time: 7232 sec

us-09-870-1113-1.rge

Tue Aug 27 07:53:39 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2002, 00:34:40 ; Search time 209.67 seconds
(without alignments)
8966.573 Million cell updates/sec

Title: US-09-870-113-1
Perfect score: 1095
Sequence: 1 atggaggttgaggcgggg.....agtggaggctgcaagtga 1095

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_032802.*	
				1:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
				2:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
				3:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
				4:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
				5:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
				6:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
				7:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
				8:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
				9:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
				10:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
				11:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
				12:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
				13:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
				14:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
				15:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
				16:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
				17:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
				18:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
				19:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
				20:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
				21:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
				22:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
				23:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
				24:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	987.4	90.2	1336	21	Human uncoupling p
2	972.4	88.8	1322	22	Human transport pr
3	958	87.5	1294	22	Human polynucleoti
4	916.5	83.7	1316	22	Human polynucleoti
5	872	79.6	1418	21	Human ORFX ORF398
6	639	58.4	710	20	Human gene express
7	576	52.6	2502	22	Human secreted pro
8	556	50.8	2037	24	cdNA sequence #574
9	520	47.5	997	22	Human mitochondria

10	508	46.4	27960	22	AAK69779	Human immune/haema
11	508	46.4	27960	22	AAK73320	Human immune/haema
12	434.2	39.7	1305	22	AAF59920	Human mitochondria
13	417.8	38.2	483	24	ABI19871	Mouse ischaemic co
14	392	35.8	401	22	AAF66156	Novel human polynu
15	380.8	34.8	1716	21	AACT77189	Human ORFX ORF2744
16	342.2	31.3	1913	22	AAH89944	Human bone marrow
17	338.4	30.9	1902	22	AAK52172	Human polynucleoti
18	330.2	29.2	2562	22	AAH90057	Human bone marrow
19	300	27.4	300	20	AAZ13026	Human gene express
20	277.2	25.3	1814	22	AAK94871	Human full-length
21	276	25.2	1015	22	AAK92501	Human mitochondria
22	276	25.2	1673	22	AAK53156	Human polynucleoti
23	276	25.2	2710	21	AAK66452	Human secreted pro
24	275.6	25.2	1890	22	AAK94897	Human full-length
25	272.2	24.9	600	22	AAH22163	Human mitochondria
26	250	22.8	8211	22	AAK80490	Human immune/haema
27	181	16.5	339	22	AAK58672	Human immune/haema
28	173.6	15.9	842	22	AAK92357	Human cdna 5'-end
29	173.6	15.9	842	22	AAK93895	Human cdna clone r
30	172.6	15.8	1468	21	AACT7175	Human ORFX ORF2730
31	172	15.7	3028	22	AAK91320	Human polynucleoti
32	171.6	15.7	2137	23	ABL08627	Drosophila melanog
33	170.4	15.6	622	22	AAK91764	Human cdna 5'-end
34	170.4	15.6	622	22	AAK93200	Human cdna clone r
35	164.2	15.0	792	21	AACT7177	Human ORFX ORF2732
36	161.4	14.7	1548	23	AAK55357	DNA encoding novel
37	160	14.6	1758	20	AAV84573	Human secreted pro
38	160	14.6	1758	22	ABA83356	Human secreted pro
39	159.6	14.6	504	23	AAK56354	DNA encoding novel
40	153.6	14.0	338	22	AAK23288	Human prostate can
41	138.2	12.6	975	21	AACT7173	Human ORFX ORF2728
42	113.4	10.4	599	22	AAK24885	Human ovarian PCr
43	113.4	10.4	599	22	AAH83525	Human ovarian tumo
44	109.8	10.0	2837	22	AAK80493	Human immune/haema
45	92	8.4	5150	18	AAK68952	Blackcurrant fruit

ALIGNMENTS

RESULT	1
ID	AAK90457 standard; cdna; 1336 BP.
XX	AAK90457;
AC	AAK90457;
XX	12-MAR-2001 (first entry)
DT	Human uncoupling protein cdna #6.
DE	Human; uncoupling protein; immunosuppressive; antiarthritic;
XX	antirheumatic; antiproliferative; cardiant; vasotropic;
KW	cerebroprotective; neuroprotective; antibacterial; ophthalmological;
KW	gastrointestinal; nephrotropic; gynaeological; vulnery; thrombolytic;
KW	gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
XX	infertility; ss.
OS	Homo sapiens.
XX	WO2000061614-A2.
PN	19-OCT-2000.
XX	06-APR-2000; 2000WO-US09534.
PF	09-APR-1999; 99US-0128701.
XX	08-JUL-1999; 99US-0142821.
PR	18-AUG-1999; 99US-0149448.
PR	12-NOV-1999; 99US-0164751.
XX	(HUMA-) HUMAN GENOME SCI INC.

DE	Human polynucleotide SEQ ID NO 1078.
XX	
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
XX	
PR	25-APR-2000; 2000US-0552317.
XX	
PR	09-JUL-2000; 2000US-0598042.
XX	
PR	19-JUL-2000; 2000US-0620312.
XX	
PR	03-AUG-2000; 2000US-0653450.
XX	
PR	14-SEP-2000; 2000US-0662191.
XX	
PR	19-OCT-2000; 2000US-0693036.
XX	
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI: 2001-442253/47.
DR	P-FSDB; AAM33719.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Claim 1; SEQ ID NO 1078; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SQ	Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;
XX	
Qy	Query Match 83.7%; Score 916.6; DB 22; Length 1316;
Db	Best Local Similarity 97.9%; Pred. No. 1.8e-195;
Qy	Matches 939; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
Db	138 cggggcctgcaggcccccgcgtacgaagatcccgactccggcccgactacgagggcgt 197
Qy	118 cggggcctgcaggcccccgcgtacgaagatcccgactccggcccgactacgagggcgt 74
Db	198 gccggctgagaccactgtccaccgacatg-gtggcaggccgctgtgacggatccctgg 256
Qy	75 gtcggctggagccactgtccaccgacatgctgggggtatgcgcgcgacatcggtctgctg 134
Db	257 agcactgcgtgatgtaccctcactgcgtcgaagaccggtacagatgcaggtctacagctg 316

Qy	136	cggggccccgccccgggctacgacacgaacccgggccccgggctacgacacgagggcct	137
Db	15	cgaggcctcgaggccccgggtacgatcagatccggactcggccgcgactacgaggcct	74
Qy	198	gcggcctggagccactgtcacccagcacatg-tggcgaggccgctggcaggatcctcgg	256
Db	75	gtcggcctggagccactgtcacccagcacatgtgggggtatgcgcgcagccatggatgctcgg	134
Qy	257	agcactcgctgattaccgccatcgactcgcgtcaagaccggatcgagagtctacagctcg	316

Db 166 gttgtcaagcagagtgatgcagtgtacaaactcaccataccaccgggtgacagactgtgta 225
Qy 637 cggcagtgtagcaaaatgaaggccggcggcctttttacccagactacacaccagctg 696
Db 226 cggcagtgtagcaaaatgaaggccggcggcctttttacccagactacacaccagctg 285
Qy 697 accatgaacgttctcttccaaagccattcacttcatgacctatgaattctcgcagagcac 756
Db 286 accatgaacgttctcttccaaagccattcacttcatgacctatgaattctcgcagagcac 345
Qy 757 tttaacccccagagcaggtacaaacccaagctccacagctctctctgagcttgcgcagga 816
Db 346 tttaacccccagagcaggtacaaacccaagctccacagctctctctgagcttgcgcagga 405
Qy 817 gctagtctccgcagcacaacccccactgacgttttgcaaaacactgctcaacaccag 876
Db 406 gctagtctccgcagcacaacccccactgacgttttgcaaaacactgctcaacaccag 465
Qy 877 gagtcttggctttgaactcacattacagagacatatcacaggcatggctagtgccttc 936
Db 466 gagtcttggctttgaactcacattacagagacatatcacaggcatggctagtgccttc 525
Qy 937 aggcaggtatatcaagttagtggtgacccgctatttccaggggtgcagccagagta 996
Db 526 aggcaggtatatcaagttagtggtgacccgctatttccaggggtgcagccagagta 585
Qy 997 attaccagatccctccacagccatcgcatggtctgtgtatgattcttcaataccta 1056
Db 586 attaccagatccctccacagccatcgcatggtctgtgtatgattcttcaataccta 645
Qy 1057 atcactaaaggcagaagtagtgaggctggcgtggcgaagtga 1095
Db 646 atcactaaaggcagaagtagtgaggctggcgtggcgaagtga 684

RESULT 7
AAS03906
ID AAS03906 standard; cDNA; 2502 BP.
XX
AC AAS03906;
XX
DT 29-AUG-2001 (first entry)
XX
DE Human secreted protein gene #25.
XX
KW Human secreted protein; autoimmune disorder; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;
KW nervous system disorder; bacterial infection; viral infection; ss;
KW fungal infection; ocular disorder; wound healing; tissue regeneration;
KW epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.
OS Homo sapiens.
XX
PN WO200123598-A1.
XX
PD 05-APR-2001.
XX
PF 26-SEP-2000; 2000WO-US26324.
XX
PR 27-SEP-1999; 99US-0155807.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Komatsoulis G, Ruben SM, Rosen CA;
XX
XX WPI; 2001-281684/29.
XX
XX P-PSDB; AAU01950, AAU01989.
XX
XX Forty one nucleic acid molecules encoding human secreted proteins, useful
XX in the prevention, treatment and diagnosis of cancer, immune disorders,
XX cardiovascular disorders and neurological diseases -
XX
XX Disclosure; Page 460-461; 518pp; English.

XX Sequences AAS03873-AAS03922 represent isolated nucleic acid molecules and
CC PCR primers of the invention. acid of the invention. Secreted proteins and
CC their related nucleic acids can be used in the diagnosis of or
CC susceptibility to a pathological condition by determining the presence or
CC absence of a mutation in a nucleic acid or the presence or amount of
CC expression of a secreted protein. The sequences are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. The antibodies to the
CC polypeptides can also be used in alleviating symptoms associated with
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). The disorders include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The peptides can also be used to aid wound healing and epithelial cell
CC proliferation, to help prevent skin ageing due to sunburn, to maintain
CC organs before transplantation, to regenerate tissues, in chemotaxis and
CC as a food additive or preservative to alter storage capabilities.
XX
SQ Sequence 2502 BP; 587 A; 655 C; 669 G; 589 T; 2 other;

Query Match 52.6%; Score 576; DB 22; Length 2502;
Best Local Similarity 100.0%; Pred. No. 2.5e-119;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 520 ggtgcggcggtgtgtggcaacattacttcatgtcagccatgaacctgcggaagtg 579
Db 1513 ggtgcggcggtgtgtggcaacattacttcatgtcagccatgaacctgcggaagtg 1572
Qy 580 gtcaagcagaggtgacagatgtacaactcaccatcaccacgggtgacagactgtgacgg 639
Db 1573 gtcaagcagaggtgacagatgtacaactcaccatcaccacgggtgacagactgtgacgg 1632
Qy 640 gcagtgtggcaaaatgaaggcgccggttttaccgcagctacacaccacagctgacc 699
Db 1633 gcagtgtggcaaaatgaaggcgccggttttaccgcagctacacaccacagctgacc 1692
Qy 700 atgaacgttcttcccaagccattcacttcatgacctatgaattcctgcaggaagcacttt 759
Db 1693 atgaacgttcttcccaagccattcacttcatgacctatgaattcctgcaggaagcacttt 1752
Qy 760 aacccccagagcgtgtacacaccccaagctccacagctcctctctgtgagcttgcgagagct 819
Db 1753 aacccccagagcgtgtacacaccccaagctccacagctcctctctgtgagcttgcgagagct 1812
Qy 820 gtactgtccgcagcacaaccccaactgacgttttgcaaaacactgctcaacaccagag 879
Db 1813 gtactgtccgcagcacaaccccaactgacgttttgcaaaacactgctcaacaccagag 1872
Qy 880 tccttggctttgaactcacacattacaggaacatatacaggaagcagctagtgcttcagg 939
Db 1873 tccttggctttgaactcacacattacaggaacatatacaggaagcagctagtgcttcagg 1932
Qy 940 acggtatatcaagtagtggtgggtgaccgcctatttccgaggggtgcagggcagagtaatt 999
Db 1933 acggtatatcaagtagtggtgggtgaccgcctatttccgaggggtgcagggcagagtaatt 1992
Qy 1000 taccagatccctccacagccatcgatggtctgtgtatgattcttcaatacctaacc 1059
Db 1993 taccagatccctccacagccatcgatggtctgtgtatgattcttcaatacctaacc 2052
Qy 1060 actaaaggcaagaagtagtgagggtggcgaagtga 1095
Db 2053 actaaaggcaagaagtagtgagggtggcgaagtga 2088

RESULT 8
AAS62787
ID AAS62787 standard; cDNA; 2037 BP.

XX AAS62787;
AC 14-FEB-2002 (first entry)
DT
DE cDNA sequence #574 encoding novel human secreted protein.
XX
XX Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW immunosuppressive; antirheumatic; ss.
XX
OS Homo sapiens.
XX
XX WO200177291-A2.
PN
XX 18-OCT-2001.
PD
XX 29-MAR-2001; 2001WO-US10485.
XX
XX 06-APR-2000; 2000US-195604P.
PR
XX (GENY) GENETICS INST INC.
PA
XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
PI
XX WPI; 2002-010900/01.
DR
XX New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease -
PT
XX
PS Claim 1; Page 364-365; 391pp; English.
XX
CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides
CC a method for producing proteins from these polynucleotide sequences.
CC The proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are
CC useful for treating diseases such as hyperproliferative disorders
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory
CC disorders (e.g. arthritis) and infectious disorders (e.g. hepatitis).
CC The polynucleotide sequences of the invention are also useful in gene
CC therapy. AAS62214-AAS62838 represent the cDNA sequences of the
CC invention that encode for novel human secreted proteins.
XX
SQ Sequence 2037 BP; 475 A; 522 C; 524 G; 515 T; 1 other;

Query Match 50.8%; Score 556; DB 24; Length 2037;
Best Local Similarity 99.0%; Pred. No. 7e-115;
Matches 570; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 520 ggtgcgcccgggtgtgtggaacattacttcatgatgcagccatgaacccctgcggaagt 579
DB 1138 gggtgcgcccgggtgtgtggaacattacttcatgatgcagccatgaacccctgcggaagt 1197
QY 580 gtcaagcagagatgcagatgtacaactaccatcaccacgggtgacagactgttacgg 639
DB 1198 gtcaagcagagatgcagatgtacaactaccatcaccacgggtgacagactgttacgg 1257
QY 640 gcagttgtgcaaaatgaaggccggggcctttaccgcagctacaccaccagctgacc 699
DB 1258 gcagttgtgcaaaatgaaggccggggcctttaccgcagctacaccaccagctgacc 1317
QY 700 atgaacgttcctttccagcagccattcacttcagctacactatgaattcctcagagacattt 759
DB 1318 atgaacgttcctttccagcagccattcacttcagctacactatgaattcctcagagacattt 1376

QY 760 aacccccagagcgtacaaacccaagctccctctctgtgagcttgcgagagct 819
DB 1377 aacccccagagcgtacaaacccaagctccctctctgtgagcgtgcgagagct 1436
QY 820 gtagtgcgcgagcacaacccccactgacgtttgcaaaacactgctcaacacccaggag 879
DB 1437 gtagtgcgcgagcacaacccccactgacgtttgcaaaacactgctcaacacccaggag 1496
QY 880 tcttggctttgaactcacacattacagagacataacagggcatggttagtccttcagg 939
DB 1497 tcttggctttgaactcacacattacagagacataacagggcatggttagtccttcagg 1556
QY 940 acggtatatcaagtgtggtgacgcctatttccagaggggtgcagccagagtaatt 999
DB 1557 acggtatatcaagtgtggtgacgcctatttccagaggggtgcagccagagtaatt 1616
QY 1000 taccagatccccccacagccatcgatggtctgtgtatgatttcttcaaatcctaattc 1059
DB 1617 taccagatccccccacagccatcgatggtctgtgtatgatttcttcaaatcctaattc 1676
QY 1060 actaaaggcaagaagatggagggtggcgaagtga 1095
DB 1677 actaaaggcaagaagatggagggtggcgaagtga 1712
RESULT 9
AAH22162
ID AAH22162 standard; cDNA; 997 BP.
XX
AC AAH22162;
XX
DT 20-AUG-2001 (first entry)
XX
DE Human mitochondrial solute carrier (hMSC-homologue) encoding cDNA SEQ.6.
XX
KW Human; hMSC; hMSC-homologue; mitochondrial solute carrier; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 379..909
FT /*tag= a
FT /product= "mitochondrial solute carrier (hMSC-homologue)"
XX
PN CN1281895-A.
XX
PD 31-JAN-2001.
XX
PF 27-JUN-2000; 2000CN-0116795.
XX
PR 27-JUN-2000; 2000CN-0116795.
XX
PA (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
PI Li N, Xu X, Xiao H;
XX
DR WPI; 2001-282781/30.
XX P-PSDB; AAB98221.
PT New human mitochondrial solute carrier (hMSC) protein isomer and its
PT code sequence -
XX
PS Claim 1; Page 17-18 (disclosure); 22pp; Chinese.
XX
CC The present invention describes a human mitochondrial solute carrier
CC protein, designated hMSC-homologue, which is expressed in human
CC hypophysis. Also described are methods for the preparation and
CC detection of hMSC-homologue protein and nucleotide sequences. The
CC present sequence encodes hMSC-homologue, as given in the present
CC invention.
XX
SQ Sequence 997 BP; 240 A; 261 C; 242 G; 254 T; 0 other;

Query Match 47.5%; Score 520; DB 22; Length 997;
Best Local Similarity 100.0%; Pred. No. 6.7e-107;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	576	agtggtcaagcagaggtgcagatgtacaactaccataccaccgggtgacagactgtgt	635
Db	390	agtggtcaagcagaggtgcagatgtacaactaccataccaccgggtgacagactgtgt	449
Qy	636	acgggcagtggtgcaaaatgaaggccggcggccttttacgcgagctacacaccacagct	695
Db	450	acgggcagtggtgcaaaatgaaggccggcggccttttacgcgagctacacaccacagct	509
Qy	696	gaccatgaacgtctctttccaagccattcacttcattgatcctatgaattcctgcaggagca	755
Db	510	gaccatgaacgtctctttccaagccattcacttcattgatcctatgaattcctgcaggagca	569
Qy	756	ctttaaccccccagagcgggtacaacccaagctcccaagctctctctgtgagcttgcgcagg	815
Db	570	ctttaaccccccagagcgggtacaacccaagctcccaagctctctctgtgagcttgcgcagg	629
Qy	816	agctgtagctccgcagcagcaacccccactggagcttgcgaacactgctcaacaccca	875
Db	630	agctgtagctccgcagcagcaacccccactggagcttgcgaacactgctcaacaccca	689
Qy	876	ggagtccttgcttgaactcacattacagacacatcacagcatggtcgtagtgcctt	935
Db	690	ggagtccttgcttgaactcacattacagacacatcacagcatggtcgtagtgcctt	749
Qy	936	caggacggtatatcaagttagtggtggtagccgctatttccgaggggtgcaggccagagt	995
Db	750	caggacggtatatcaagttagtggtggtagccgctatttccgaggggtgcaggccagagt	809
Qy	996	aattaccagatccctccacagccatcgcatgctgtgtatgattcttcaataacct	1055
Db	810	aattaccagatccctccacagccatcgcatgctgtgtatgattcttcaataacct	869
Qy	1056	aatcactaaaagcacaagaagtgaggctggcgaagtga	1095
Db	870	aatcactaaaagcacaagaagtgaggctggcgaagtga	909
RESULT 10			
AAK69779			
ID	AAK69779 standard; DNA; 27960 BP.		
XX	AC AAK69779;		
XX	06-NOV-2001 (first entry)		
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24591.		
XX	Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;		
KW	cytostatic; gene therapy; vaccine; metastasis; ds.		
XX	Homo sapiens.		
OS	WO200157182-A2.		
PN	09-AUG-2001.		
PD	17-JAN-2001; 2001WO-US01354.		
PF	31-JAN-2000; 2000US-0179065.		
XX	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		
PR	14-AUG-2000; 2000US-0225758.		
PR	14-AUG-2000; 2000US-0225759.		
PR	18-AUG-2000; 2000US-0226279.		
PR	22-AUG-2000; 2000US-0226681.		
PR	22-AUG-2000; 2000US-0226868.		
PR	22-AUG-2000; 2000US-0227182.		
PR	23-AUG-2000; 2000US-0227009.		
PR	30-AUG-2000; 2000US-0228924.		
PR	01-SEP-2000; 2000US-0229287.		
PR	01-SEP-2000; 2000US-0229343.		
PR	01-SEP-2000; 2000US-0229344.		
PR	01-SEP-2000; 2000US-0229345.		
PR	05-SEP-2000; 2000US-0229509.		
PR	05-SEP-2000; 2000US-0229513.		
PR	06-SEP-2000; 2000US-0230437.		
PR	06-SEP-2000; 2000US-0230438.		
PR	08-SEP-2000; 2000US-0231242.		
PR	08-SEP-2000; 2000US-0231243.		
PR	08-SEP-2000; 2000US-0231244.		
PR	08-SEP-2000; 2000US-0231413.		
PR	08-SEP-2000; 2000US-0231414.		
PR	08-SEP-2000; 2000US-0232081.		
PR	12-SEP-2000; 2000US-0231968.		
PR	14-SEP-2000; 2000US-0232397.		
PR	14-SEP-2000; 2000US-0232398.		
PR	14-SEP-2000; 2000US-0232399.		
PR	14-SEP-2000; 2000US-0232400.		
PR	14-SEP-2000; 2000US-0232401.		
PR	14-SEP-2000; 2000US-0233063.		
PR	14-SEP-2000; 2000US-0233064.		
PR	14-SEP-2000; 2000US-0233065.		
PR	21-SEP-2000; 2000US-0234223.		
PR	21-SEP-2000; 2000US-0234274.		
PR	25-SEP-2000; 2000US-0234997.		
PR	25-SEP-2000; 2000US-0234998.		
PR	26-SEP-2000; 2000US-0235484.		
PR	27-SEP-2000; 2000US-0235834.		
PR	27-SEP-2000; 2000US-0235836.		
PR	29-SEP-2000; 2000US-0236327.		
PR	29-SEP-2000; 2000US-0236367.		
PR	29-SEP-2000; 2000US-0236368.		
PR	29-SEP-2000; 2000US-0236369.		
PR	29-SEP-2000; 2000US-0236370.		
PR	02-OCT-2000; 2000US-0236802.		
PR	02-OCT-2000; 2000US-0237037.		
PR	02-OCT-2000; 2000US-0237038.		
PR	02-OCT-2000; 2000US-0237039.		
PR	02-OCT-2000; 2000US-0237040.		
PR	13-OCT-2000; 2000US-0239935.		
PR	13-OCT-2000; 2000US-0239937.		
PR	20-OCT-2000; 2000US-0240960.		
PR	20-OCT-2000; 2000US-0241221.		
PR	20-OCT-2000; 2000US-0241785.		

20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 24591; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;
SQ
Query Match 46.4%; Score 508; DB 22; Length 27960;
Best Local Similarity 99.8%; Pred. No. 6.3e-104;
Matches 519; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 576 agtgggtcaagcagagatgcagatgtacaaactcacatcacaccgggtgacagactgtgt 635
|||||
Db 20845 agtgggtcaagcagagatgcagatgtacaaactcacatcacaccgggtgacagactgtgt 20904
QY 636 acgggcagtggtgcaaatgaaggccggggcctttaccgcagctacacccacagct 695
|||||
Db 20905 acgggcagtggtgcaaatgaaggccggggcctttaccgcagctacacccacagct 20963
QY 696 gaccatgaacgttcttccaaagcattcattcagctacacattctctgcaggagca 755
|||||
Db 20964 gaccatgaacgttcttccaaagcattcattcagctacacattctctgcaggagca 21023
QY 756 cttaaacccccagagcgtgtacaaacccaaagctccacgtctctctgagcttgcgcagg 815
|||||
Db 21024 cttaaacccccagagcgtgtacaaacccaaagctccacgtctctctgagcttgcgcagg 21083
QY 816 agctgtagctgccgcagccacaaaccccaactggacgtttgcaaacactgtctcaacacca 875
|||||
Db 21084 agctgtagctgccgcagccacaaaccccaactggacgtttgcaaacactgtctcaacacca 21143
QY 876 ggagctcttggtttgaactcacacattacagacacatacacaggcagctagtcctt 935
|||||
Db 21144 ggagctcttggtttgaactcacacattacagacacatacacaggcagctagtcctt 21203
QY 936 caggacggttatatacaagtagtggtgggtgacccgctatttccgaggggtgcagccagagt 995
|||||
Db 21204 caggacggttatatacaagtagtggtgggtgacccgctatttccgaggggtgcagccagagt 21263
QY 996 aattaccagatccctccacagccatcgcatggtctgtgtatgagttcttcaaatacct 1055
|||||
Db 21264 aattaccagatccctccacagccatcgcatggtctgtgtatgagttcttcaaatacct 21323
QY 1056 aatcactaaaaggcaagaagagtgagggtggcctggcaagtga 1095
|||||
Db 21324 aatcactaaaaggcaagaagagtgagggtggcctggcaagtga 21363
RESULT 11
AAK73320
ID AAK73320 standard; DNA; 27960 BP.
XX
AC AAK73320;
XX
XX 06-NOV-2001 (first entry)
XX
DE Human Immune/hematopoietic antigen genomic sequence SEQ ID NO:28132.
XX
XX Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX

PD 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
DR


```

XX (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
XX WPI: 2001-091805/10.
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 818; 1046pp; English.
XX
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match      35.8%; Score 392; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.2e-78;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 gtctacagcctgaccagctgcccgctatgcgaatgtgttgaggcccttgaggatta 364
Db 9 gtctacagcctgaccagctgcccgctatgcgaatgtgttgaggcccttgaggatta 68
QY 365 taagaacggaggcctatgagcccataggggctgaacgtcacagcaacaggcgag 424
Db 69 taagaacggaggcctatgagggcccataggggctgaacgtcacagcaacaggcgag 128
QY 425 ggccgtgccacgcctttatttgcctgctacgaaagttaaaagacattgagtgatg 484
Db 129 ggccgtgccacgcctttatttgcctgctacgaaagttaaaagacattgagtgatg 188
QY 485 taatcacacctggggcaatagcattatgcgaatgtgtgcggccggtgtgtggaacat 544
Db 189 taatcacacctggggcaatagcattatgcgaatgtgtgcggccggtgtgtggaacat 248
QY 545 tacttcatgatgcagcattgaacctgcggaagtgttcaagcagagatgcagattgata 604
Db 249 tacttcatgatgcagcattgaacctgcggaagtgttcaagcagagatgcagattgata 308
QY 605 actcacataccaccgggtgacagatgtgtgtacggcgagtggtggcaaatgaaggggccg 664
Db 309 actcacataccaccgggtgacagatgtgtgtacggcgagtggtggcaaatgaaggggccg 368
QY 665 gggccctttaccagctacacaccagctg 696
Db 369 gggccctttaccagctacacaccagctg 400

RESULT 15
AAC77189
ID AAC77189 standard; cDNA; 1716 BP.
XX

```

```

AC AAC77189;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2744 polynucleotide sequence SEQ ID NO:5487.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
WPI: 2000-602362/57.
DR P-PSDB; AAB42980.
DR
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
PT
PS Claim 5; Page 4661-4662; 5507pp; English.
XX
CC AAC7446 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 1716 BP; 411 A; 457 C; 446 G; 402 T; 0 other;

```

Query Match 34.8%; Score 380.8; DB 21; Length 1716;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2002, 01:24:15 ; Search time 49.51 Seconds
(without alignments)
5432.613 Million cell updates/sec

Title: US-09-870-113-1

Perfect score: 1095

Sequence: 1 atggagtgaggcgaggcgagg.....agtggaggtggcgcaagtga 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	8.4	5150	4	US-09-068-140A-14
2	90.4	8.3	1311	4	US-09-068-140A-9
3	62.4	5.7	319	4	US-09-165-264-8
4	58.2	5.3	318	4	US-09-165-264-12
5	58.2	5.3	320	4	US-09-165-264-14
6	58.2	5.3	152331	3	US-09-128-155-16
7	57.8	5.3	320	4	US-09-165-264-7
8	57.6	5.3	320	4	US-09-165-264-13
9	57.4	5.2	320	4	US-09-165-264-11
10	50.6	4.6	833	2	US-08-403-852B-3
11	50.6	4.6	833	3	US-08-510-646B-3
12	50.6	4.6	833	4	US-09-231-818-3
13	50.6	4.6	5392	2	US-08-403-852B-1
14	50.6	4.6	5392	3	US-08-510-646B-1
15	50.6	4.6	5392	4	US-09-231-818-1
16	49.4	4.5	43280	2	US-08-804-227C-1
17	47	4.3	1227	4	US-08-903-800A-1
18	46.8	4.3	1931	2	US-09-130-114-2
19	46.8	4.3	4257	2	US-08-690-473-1
20	46.8	4.3	4257	4	US-09-259-821A-1
21	46.8	4.3	4257	1	US-08-843-659-1
22	46.8	4.3	12001	1	US-08-458-568A-11
23	46.6	4.3	4403765	4	US-09-103-840A-2
24	46.6	4.3	4411529	4	US-09-103-840A-1
25	46.2	4.2	2823	1	US-08-398-008A-1
26	46.2	4.2	2823	2	US-08-893-333-1
27	46.2	4.2	3415	1	US-08-054-077C-1

28	45.8	4.2	460	2	US-08-903-800A-4
29	45.6	4.2	1300	4	US-08-483-533-39
30	45.6	4.2	1300	4	US-09-283-471A-39
31	45.6	4.2	53526	3	US-08-658-136-2
32	45.6	4.2	53577	3	US-08-658-136-1
33	45.4	4.1	1327	4	US-08-483-533-36
34	45.4	4.1	1327	4	US-09-283-471A-36
35	45.4	4.1	1335	5	PCT-US91-065332-1
36	45.2	4.1	1028	4	US-08-118-200-1
37	45.2	4.1	1028	4	US-08-458-745-1
38	45.2	4.1	12001	1	US-08-458-568A-11
39	45	4.1	8438	1	US-07-945-283-1
40	44.6	4.1	5228	4	US-09-428-711A-15
41	44.6	4.1	23673	4	US-09-773-816-1
42	44.4	4.1	405	2	US-08-903-800A-3
43	44.2	4.0	700	4	US-09-236-097-7
44	44.2	4.0	1026	4	US-07-751-891B-24
45	44.2	4.0	1868	1	US-08-658-883B-1

ALIGNMENTS

RESULT 1
US-09-068-140A-14
; Sequence 14, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-14

```
Query Match      8.4%; Score 92; DB 4; Length 5150;
Best Local Similarity 49.3%; Pred. No. 3.4e-13;
Matches 347; Conservative 0; Mismatches 330; Indels 27; Gaps 3;

Qy 189 cgaggcgtgcggctgagaccactgtcacacgcaacatggtggcaggcgcggtggcagg 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3236 CGCGCGGTGCGATGATGCGCTTCACTTGGCAGTTCATGATGCGCGGTTCATCGCTGG 3295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 249 gatcctgagcactgctgagtacccatgactgctcaagacccgagtgacagtgct 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3296 ATCAATCGAGCATATGCGGATGATCCGGTGTGATACGCTTAAACTCCGATACAGGGTAT 3355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 309 acagcctgaccagctgcccgtatcgcaatggttggaggccctctggagattataag 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3356 TGGGTTCATGTTGGCT---CAATCCGCGGTCTCCGACAAGCCCTTGGGTGCTGACTGAA 3412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 369 aacgagggcctatgagggcccatgaggggctgaacgtcacagcaacaggcgcaggcc 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3413 AGTTGAAGGTCCCGCGGACTTTACCGTGGCATTTGGTGAATGGTCTCGGTGCGAGGACC 3472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 429 tcccacgccccttatttgcctgctacgaaagttaaagaacacatgagtgtgta 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3473 AGCTCAGGAGTGTATTTCTCGTTTACGATGTGTAAAGAGACTTTTCTCATGGTGA 3532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 489 ccaccctggggcaatagccatattgccaatggtgcggcggtgtgtggcaacattact 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3533 TCCGAGCAATTCGGTGGCGACGCGTTTCGGGGGTGTTCCGGACGGTGGCAA----- 3585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 549 tcattgatcagccatgaaccctgcgaagtggtcaagcagagagatgcagatgataactc 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3586 --GCGACGCGGTGATACCGCGATGGATGTGGTGAACAGAGGTTGCGAGTTGCAGAGCAG 3643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 609 accataccacgggtgacagactgtgacggcagtggtggcaaatgaggggcccgggc 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3644 TCCGTACAGGGTGTGTGATGGCTGAGGAGGTGTGGTAGAAGAGGATTTGGCGC 3703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 669 cttttaccgcagctacacacccagctgaccatgaacgtctcttccaaagccattcact 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3704 ATTTTACGCATCTTATCGAACAACCTGTGTGTCATGAATGCCCGGTTTACGGCGGTTCACTT 3763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 729 catgacctatgaattcctgcagagacactttaacccccagagacggttacaccccaagctc 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3764 CCCCACATATGAAGCCAGCAAGAAAGGTTGTGGAGGTGTGCGCGGAGACTGCGCAAGCA 3823
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 789 ccagctctct-----ctggagcttgcagagagctgtagctgcgcagc 833
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3824 TGAGAAATTTGTAGTGCATGTACTGTGCTGGTCTGCTGTGGAGCTTTGGCTGCAGTAGT 3883
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 834 cacaacccactggagcttgcataaacactgctcaacacccagg 877
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3884 AACCACTCCACTAGATGTGTCAAAACTCAGTTGTCAGTGCAGTCCCAAG 3927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-09-068-140A-9
; Sequence 9, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068,140A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/04807
FILING DATE: No. 6281409ember 4, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L., 680
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: C70237
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Ribes nigrum
STRAIN: Ben Alder
US-09-068-140A-9

Query Match      8.3%; Score 90.4; DB 4; Length 1311;
Best Local Similarity 49.1%; Pred. No. 5.7e-13;
Matches 346; Conservative 0; Mismatches 331; Indels 27; Gaps 3;

Qy 189 cgaggcgtgcggctgagaccactgtcacacgcaacatggtggcaggcgcggtggcagg 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72 CGCGCGGTGCGATGATGCGCTTCACTTGGCAGTTCATGATGCGCGGTTCATCGCTGG 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 249 gatcctgagcactgctgagtacccatgactgctcaagacccgagtgacagtgct 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 ATCAATCGAGCATATGCGGATGATCCGGTGTGATACGCTTAAACTCCGATACAGGCTAT 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 309 acagcctgaccagctgcccgtatcgcaatggttggaggccctctggagattataag 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 TGGGTTCATGTTGGCT---CAATCCGCGGTCTCCGACAAGCCCTTGGGTGCTGATACTGAA 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 369 aacgagggcctatgagggcccatgaggggctgaacgtcacagcaacaggcgcaggcc 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 AGTTGAAGGTCCCGCGGACTTTACCGTGGCATTTGGTGAATGGTCTCGGTGCGAGGACC 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 429 tcccacgccccttatttgcctgctacgaaagttaaagaacacatgagtgtgta 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 AGCTCAGCAGTGTATTTCTCGTTTACGAGATGTGTAAAGAGACTTTTCTCATGGTGA 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 489 ccaccctggggcaatagccatattgccaatggtgcggcggtgtgtggcaacattact 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 TCCGAGCAATTCGGTGGCGACGCGTTTCGGGGGTGTTCCGGAGGTTGGCAA----- 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 549 tcattgatcagccatgaaccctgcgaagtggtcaagcagagagatgcagatgataactc 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 --GCGACGCGGTGATACCGCGATGGATGTGGTGAACAGAGGTTGCTAGTTCAGAGCAG 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 609 accataccacgggtgacagactgtgacggcagtggtggcaaatgaggggcccgggc 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 TCCGTACAAAGGTTGTGTGATTTGCGTGAGGAGGTGTGTGTAGAAGAGGATTTGGCGC 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 669 cttttaccgcagctacacacccagctgaccatgaacgtctcttccaaagccattcact 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 ATTTTACGCATCTTATCGAACAACCTGTGTGTCATGAATGCCCGGTTTACGGCGGTTCACTT 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 729 catgacctatgaattcctgcagagacactttaacccccagagacggttacaccccaagctc 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-0128-155-16

```

[illegible]

```

RESULT 7
US-09-165-264-7
; Sequence 7, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamcorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer sequence
US-09-165-264-7

```

[illegible]

```

RESULT      8
US-09-165-264-13
; Sequence 13, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinavagamoorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence
US-09-165-264-13

```

[illegible]

```

RESULT          9
US-09-165-264-11
; Sequence 11, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamoorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-11

```

	Query Match	5.2%	Score 57.4	DB 4	Length 320
	Best Local Similarity	61.9%	Pred. No. 3.e-05		
	Matches 91	Conservative	0	Mismatches 56	Indels 0
Qy	3	ggaggttgaggggcgggtgctgctggcggtgtgctggcgggggcgccggcgagggcccgggcg	62		
Db	158	ggg	217		
Qy	63	gagcccggggagtcggcgctctctggacggttgagctgacggcgctggtctgacggcgggggc	122		
Db	218	ggag	277		

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2002, 00:28:35 ; Search time 1719.3 Seconds

(without alignments)

8596.037 Million cell updates/sec

Title: US-09-870-113-1

Perfect score: 1095

Sequence: 1 atggagttggagggcgagg.....agtggagggctggcaagtga 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	692.2	63.2	888	9	AL530804
2	637.4	58.2	747	10	BI827322
3	615.8	56.2	1131	10	BG295496
4	601.4	54.9	681	9	AV704087
5	587	53.6	826	10	BI254253
6	586	53.5	986	10	BG331197
7	560.4	51.2	645	10	BG964218
8	543	49.6	855	10	BI412175
9	534.6	48.8	817	10	BI103329
10	532.2	48.6	576	10	BI057218
11	517	47.2	927	10	BG972351
12	507.6	46.4	574	10	BI118580
13	507.4	46.3	908	10	BG697119
14	502.4	45.9	560	10	BI794616
15	502.2	45.9	598	10	BE913718
16	473	43.2	808	10	BG918082
17	472.6	43.2	531	9	AW915209

18	470.6	43.0	563	10	BF193023
19	468	42.7	581	9	AW341177
20	453.6	41.4	830	10	BF143133
21	453.4	41.4	578	10	BM483098
22	441.4	40.3	480	10	BI045863
23	436	39.8	1295	11	AK015770
24	435	39.7	451	10	BF515180
25	433.8	39.6	529	9	AA431276
26	433	39.5	501	9	AW211366
27	433	39.5	527	9	AI797090
28	432	39.5	526	9	AI797100
29	430.2	39.3	1209	11	AK006155
30	429.6	39.2	741	9	AL563845
31	427.2	39.0	709	9	AJ397503
32	419.4	38.3	573	10	BI541703
33	417.8	38.2	483	9	AA061624
34	409.8	37.4	697	10	BI854643
35	405.6	37.0	563	10	BF078769
36	398.4	36.4	541	9	AA234031
37	398.2	36.4	967	10	BI739710
38	393.6	35.9	453	10	BF916224
39	391	35.7	488	9	AI797101
40	387	35.3	718	10	BJ041453
41	383.2	35.0	423	10	BI339456
42	382.8	35.0	425	10	BI304641
43	381.6	34.8	848	9	AI743110
44	371	33.9	434	10	W53634
45	368.2	33.6	491	10	BF076673

ALIGNMENTS

RESULT 1

AL530804

LOCUS

DEFINITION

AL530804 LTI_NFL001_NBC4 Homo sapiens

prime, mRNA sequence.

ACCESSION

AL530804

VERSION

AL530804.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 888)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

JOURNAL

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="LTI_NFL001_NBC4"

/sex="male"

/tissue.type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 168 a 244 c 322 g 139 t 15 others

AL530804 888 bp mRNA linear EST 13-FEB-2001

AL530804 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD008YH17 5

prime, mRNA sequence.

AL530804

AL530804.1

GI:12794297

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 888)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="CS0DD008YH17"

/sex="male"

/tissue.type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 168 a 244 c 322 g 139 t 15 others

FEATURES	This clone is available at CHGC in Shanghai.	RESULT	5
	Location/Qualifiers	BI254253	
	1..681	LOCUS	
	/organism="Homo sapiens"	DEFINITION	
	/db_xref="taxon:9606"	602974729Fl NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114190 5',	
	/clone="ADBAFE05"	BI254253	826 bp mRNA linear EST 17-JUL-2001
	/clone_lib="ADS"	mrna sequence.	
	/tissue_type="Adrenal gland"	ACCESSION	BI254253
	/dev_stage="Adult"	VERSION	BI254253.1 GI:14806485
	/lab_host="SOLR"	KEYWORDS	EST.
	SOURCE	human.	
	ORGANISM	Homo sapiens	

ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 826)
AUTHORS	NH-GMC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-i@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11277 row: i column: 07
High quality sequence stop: 825.

cDNA Library Preparation: Inyte Technologies, Inc.
 cDNA Library Arrayed by: Inyte Genomics, Inc.
 DNA Sequencing by: Inyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11277 row: i column: 07
 High quality sequence stop: 825.

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11277 row: 1 column: 07
High quality sequence stop: 825.
Location/Qualifiers
1 826

```

ncftp://image:1234567890
Plate: LLAM11277 row: 1 column: 07
High quality sequence stop: 825.

FEATURES
      source
      1. .826
      /organism="Homo sapiens"
      /db="ncbi"
      /dbxref="NC_005068"

```

```

FEATURES
  source
    1. .826
      Location/Qualifiers
        High quality sequence stop: 825.
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5114190"

```

```
source
1. .826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5114190"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
```

```

/db_xref="taxon:9606"
/clone="IMAGE:5114190"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1; No

```

```

/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: Not
Site_2: SalI; Cloned unidirectionally. Primer: Oligo
average insert size 1.4 kb. Library prepared by Life

```

/tissue_type="cervical carcinoma cell line
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMW-SPORT6; Site.1: Not
 Site.2: SalI; Cloned unidirectionally. Primer: Oligo
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

BASE COUNT	211 a	233 C	208 g	174 t
ORIGIN				

BASE COUNT	Average insert size 1.4 kb. Library prepared by Life Technologies. ^a		
ORIGIN	211 a	233 c	174 t
1	100	100	100
2	100	100	100
3	100	100	100
4	100	100	100
5	100	100	100
6	100	100	100
7	100	100	100
8	100	100	100
9	100	100	100
10	100	100	100
11	100	100	100
12	100	100	100
13	100	100	100
14	100	100	100
15	100	100	100
16	100	100	100
17	100	100	100
18	100	100	100
19	100	100	100
20	100	100	100
21	100	100	100
22	100	100	100
23	100	100	100
24	100	100	100
25	100	100	100
26	100	100	100
27	100	100	100
28	100	100	100
29	100	100	100
30	100	100	100
31	100	100	100
32	100	100	100
33	100	100	100
34	100	100	100
35	100	100	100
36	100	100	100
37	100	100	100
38	100	100	100
39	100	100	100
40	100	100	100
41	100	100	100
42	100	100	100
43	100	100	100
44	100	100	100
45	100	100	100
46	100	100	100
47	100	100	100
48	100	100	100
49	100	100	100
50	100	100	100
51	100	100	100
52	100	100	100
53	100	100	100
54	100	100	100
55	100	100	100
56	100	100	100
57	100	100	100
58	100	100	100
59	100	100	100
60	100	100	100
61	100	100	100
62	100	100	100
63	100	100	100
64	100	100	100
65	100	100	100
66	100	100	100
67	100	100	100
68	100	100	100
69	100	100	100
70	100	100	100
71	100	100	100
72	100	100	100
73	100	100	100
74	100	100	100
75	100	100	100
76	100	100	100
77	100	100	100
78	100	100	100
79	100	100	100
80	100	100	100
81	100	100	100
82	100	100	100
83	100	100	100
84	100	100	100
85	100	100	100
86	100	100	100
87	100	100	100
88	100	100	100
89	100	100	100
90	100	100	100
91	100	100	100
92	100	100	100
93	100	100	100
94	100	100	100
95	100	100	100
96			

BASE COUNT	211 a	233 c	208 g	174 t
ORIGIN				
Query Match		53.6%	Score 587;	DB:10; Length 836;
Best Local Similarity		90.9%	Prod No. 5	2e-107.

Query Match 53.6%; Score 587; DB:10; Length 826;
Best Local Similarity 99.8%; Pred. No. 5.2e-107;
Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps

Query Match 53.6%; Score 587; DB:10; Length 836;
Best Local Similarity 99.88; Pred. No. 5.2e-107;
Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps

Qy 497 gggccaatagccattatgccaatgtgcgcgggtgtgtggcaacattactcatgatg 556
 Best Local Similarity 99.8%; Pred. No. 5.2e-107;
 Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps

Db 1 GGGGCATAGCCATA-TGCCAATGTGCGCCGGGTGTGGCAACATTACTTCATGATG 59

Qy 497 ggggcaatagccattattgccaatggtggccgggggtgtgtggcaacattacttcatgatg 556
Db 1 GGGCAATAGCCATA-TGCCAATGGTGGCCGGGGTGTGGCAACATTACTTTCATGATG 59
557 aagccatcaacccctggcgaactatcgaagcaaggaatgcagatgtatcaaacatcaccatacc 616

Qy	557	cagccatgaacccctgcggaagtgggtcaagcagaggatgcagatgtacaactaccattacc	616
Db	1	GGGCAATAGCCATA-IGCCAATGGTCGGCGCGGGTGTGTGGCAACATTACTTTCATGATG	59

DB I GGGGCAATAGCCATA-TGCCAATGGTGGGCGGGGCTGGGCAACAATTAATTCATCAATGAAG 616

Qy 557 cagccatgaacctgcggaagtggccaagcagaggatgcagatgtacatactaccataacc 616
|||||

DB 60 CAGCCATGAACCTTGGCGAAGTGGTCAAGCAGAGGATGCAGATGTACAACCTCACCATAAC 119
|||||

Qy	557	cagccatgaacccctgcggaagtgtatcaagcagagatgcagatgtacaactcaccataacc	616
Db	60	CAGCCATGAACCCCTGCGGAAGTGGTCAAGCAGAGATGCAGATGTACAACCTACCATAACC	119
Qy	617	accgggtgacagactgtgtacggggcagctgtggcaaaatgaaggggcccggggccctttttacc	676

Db	60	CAGCCATCAACCTTGGCGGAAGTGGTCAAGCAGAGGATGCAGATGTACAACTCACCATAACC	119
Qy	617	accgggtgcagagactgtgtacgggcagctgtggcaaaatgaaggggccggggccctttacc	676
Db	120	ACCGGGTGACAGACTGTGTACGGGCAGCTGTGGCAAAATGAAGGGGCCGGGGCCCTTTTACC	179

Qy	617	accgggtgacagactgtgtacgggcagctgtgcaaatgaaggcgccggcccttttaac	676
Db	120	ACCGGGTGACAGACTGTGTACGGGCAGTGTGGCAAAATGAAGGGCGCGGGCCCTTTTACC	179
Qy	677	gcagctacaccacccagctgaacatgaacgttctcttccaaagccattcaccttcacct	736

Db	120	ACCGGGTGACAGACTGTGTACGGGCACTGTGGCAAAATGAAGGGGGCGGGCCCTTTTACC	179
Qy	677	gcagctaccacccagctgacacatgaactgcctttccaagccattcacttcatacact	736
Db	180	GCAGCTACACACCCAGCTGACCATGAACCTGCTTCTCCAAAGGCATTCACCTTCATGACCT	239

	Qy	Db	Qy
677	gcagctacacacccagctgacctgaacgtctcttccaagcattcacttcacgt	736	
180	gcagctacacacccagctgacctgaacgtctcttccaagcattcacttcacgt	239	
737	atgaattcctgcaggagcatttaacccccagagcgttacaaccacagctccacgctcc	796	

Db	180	GCACCTACACCCAGCTGACCATGACGTTCTTTCAGGCATTCATTATGACCT	2391
Qy	737	atgaattcctgcagagcactttaacccccagagcgttacaacccaagctcccaagtc	796
Db	240	ATGAATTCCTGCAGAGCAGCTTAACCCCCAGAGAGCGTACACCCAAGCTCCACGTC	2999

[illegible]

QY 73/ atgaattccctgcaggagcagcacttttaacccccccagagagcgtacaacccaagctcccaagctcc
 |||||
 Db 240 ATGAATTCTCGAGGAGCACTTTTAACCCCCCAGAGACGGTACAACCCAAGCTCCACAGTCC 299

Qy	797	tctctggagcttgcgcagagagctagctgcccgcagccacacccccactgacgtttgca	856
Db	300	TCTCTGGAGCTTGGCGAGGAGCTAGCTGCGCAGGCACACAACCCCACTGGACGTTTGA	359
Qy	857	aaacactgctcaacacccagagtccttggcttgaactcaacattacagacatatca	916
Db	360	AAACACTGCTCAACACCCAGAGGTCCTTGGCTTTGAACTCACACATTACAGGACATATCA	419
Qy	917	caggcatgctagtgcttcaggcgggtatatacagtagtggtgggtgacccctatttcc	976
Db	420	CAGGCATGGCTAGTGCCTTTCAGGACGCTATATCAAGTAGGTGGGTGACCGCCTATTTC	479
Qy	977	gaggggtgcagggcagagtagtaattaccagatccccctccacagccatgcgtgtctgtg	1036
Db	480	GAGGGGTGCAGGCGCAGAGTAATTTACCAAGATCCCTCCACAGCCATCGCATGCTCTGT	539
Qy	1037	atgagcttctcaataactcaactcaataaagcgaagagagtagtgagggctgcgaagga	1095
Db	540	ATGAGTCTTCAATACCTTAATCACTAAAGGCAAGAGTAGTGGGGCTGGCAAGTGA	598
RESULT 6			
LOCUS BG331197 986 bp mRNA linear EST 27-FEB-2001			
DEFINITION 602431907F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4549801 5',			
mRNA sequence.			
ACCESSION BG331197			
VERSION BG331197.1 GI:13137635			
KEYWORDS EST.			
SOURCE human.			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1 (bases 1 to 986)			
AUTHORS NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL Unpublished (1999)			
COMMENT Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-r@mail.nih.gov			
Tissue Procurement: DCTD/FTP/Gazdar			
cDNA Library Preparation: Ling Hong/Rubin Laboratory			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLCM1241 row: m column: 02			
High quality sequence stop: 587.			
FEATURES			
source			
1..986			
Location/Qualifiers			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:4549801"			
/clone_lib="NIH_MGC_18"			
/tissue_type="large cell carcinoma"			
/lab_host="DH10B (phage-resistant)"			
/notes="Organ: lung; Vector: pOTB7; Site: 1; XhoI: Site 2;			
EcoRI; cDNA made by oligo-dT priming. Directionally cloned			
into EcoRI/XhoI sites using the following 5' adaptor:			
GGCCAGAG(G). Library constructed by Ling Hong in the			
laboratory of Gerald M. Rubin (University of California,			
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and			
Superscript II RT (Life Technologies). Note: this is a			
NIH_MGC Library."			
BASE COUNT 268 a 277 c 248 g 193 t			
ORIGIN			
Query Match 53.5%; Score 586; DB 10; Length 986;			
Best Local Similarity 97.5%; Pred. No. 8.5e-107;			
Matches 595; Conservative 0; Mismatches 15; Indels 0; Gaps 0;			

FEATURES		Location/Qualifiers										
source		1..645										
		/organism="Mus musculus"										
		/strain="FVB/N"										
		/db_xref="taxon:10090"										
		/clone="IMAGE:4983815"										
		/clone_lib="NCI_CGAP_Co24"										
		/lab_host="DHI0B (T1 phage-resistant)"										
		/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."										
BASE COUNT	147 a	189 c	193 g	116 t								
ORIGIN												
Query Match												
Best Local Similarity 51.2%; Score 560.4; DB 10; Length 645;												
Matches 591; Conservative 0; Mismatches 51; Indels 0; Gaps 0;												
QY	129	cggggagcgccggtcagggcccccgtacgacaagatccggactccggcccgacta	188									
Db	4	CGCGTACGGGGGGCTATPACGCCCTGTACGGCTGGATCGGAGTCGGGCCGGA	63									
QY	189	cgaggcgctgcggctggagccactgtcaccacgacatggtgaggcgccgtggcagg	248									
Db	64	CGAAGCGCTGCCGGTGGAGCCACTGTACACGACATGTTGGGGCGCGTGGCAGG	123									
QY	249	gatcctggagcactggtgatgatacccatcactgcgtcgaagaccgagtcagagtct	308									
Db	124	GATCTGGAGCATTCGGTGATGATACCCGATCGACTGCGTCAAGACCCGGATCGAGGCT	183									
QY	309	acagcctgaccagctgcccgtatcgcaatgtgttgaggccctctggaggattataag	368									
Db	184	ACAGCGTGACCCAGCGCCCGCTATCGGAACCTGTTGGAGGCTCTCTGGAGAATCATGAG	243									
QY	369	aacgagggcctatgaggcccatgaggggctgaacgtcacaacagcgagcgaggcc	428									
Db	244	GACAGAGGCCCTGTGGAGGCCCATCGGGGGCTGAACGTCACAGCAACAGCGCGGGGCC	303									
QY	429	tgccacgcctttatttgcctgtacgaaagtttaaaagacattgagtgtgtaatt	488									
Db	304	TGCCACAGCCCTCTATTTTGGCTGTACGAAAGTTTAAAAAGACATTTGAGTACGTAA	363									
QY	489	ccaccttgggggaatagccatattgccaatggtggtggccgggtgtgtggcaacattact	548									
Db	364	CCACCCAGGGGCAATAGCCATATTGCCAATGTTGCCAGCCGGATGTGTGCCACATTACT	423									
QY	549	tcatacagccatgaaccctgcggaagtgtgtcaagcagaggtgcagatgtacaactc	608									
Db	424	TATGATGAGCCATGAATCCAGCGGAAGTCGTCAAGCAGAGGATGCGAGATGTACAATC	483									
QY	609	acctaaccacgggtgacagactgtgtacggcgagtggtggcaaatgaaggcgaggcc	668									
Db	484	GCCGTACCCCGGTGACAGACTGTGTTGGGCAGTGTGGCAAAATGAAGGGGCCGGGGC	543									
QY	669	cttttacgcgagtacacccagctgaccatgaacgttcttcccaagccattcactt	728									
Db	544	CTTTTACCGCAGCTACACAGCCAGCTGACCATGAATGTCTCCCTTCCAAAGCCATTCATT	603									
QY	729	catgacctatgaattcctgcaggagcactttaacccccagag	770									
Db	604	CATGACCTATGAGTTCTTGCAGAGCACTTTAAGCCCCCAGAG	645									
RESULT 8												
BI412175												
LOCUS												
DEFINITION 602965377F1 NCI_CGAP_Lu33 Mus musculus cdna clone IMAGE:5120755 5',												
mRNA sequence.												
ACCESSION BI412175												
VERSION BI412175.1 GI:15173098												
KEYWORDS EST.												

SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 855)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-re@mail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cdna Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11294 row: j column: 20 High quality sequence start: 7 High quality sequence stop: 809. Location/Qualifiers 1..855 /organism="Mus musculus" /strain="CZECH II" /db_xref="taxon:10090" /clone="IMAGE:5120755" /clone_lib="NCI_CGAP_Lu33" /tissue_type="pooled lung tumors" /lab_host="DHI0B (phage-resistant)" /note="organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cdna was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCTCTCTTTTTTTTTTTT 3']. Double-stranded cdna was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT	209 a 233 c 241 g 172 t
ORIGIN	
Query Match	49.6%; Score 543; DB 10; Length 855;
Best Local Similarity	88.2%; Pred. No. 3e-98;
Matches 738; Conservative	0; Mismatches 85; Indels 14; Gaps 13;
QY	229 gtggcaggcgccgtggcaggatcctggagcactgcgtgtatcccccatcgactgcgtc 288
Db	18 GTGGCGGGCGCGTGGCA-GGATCCTGGAGCATTCGCTGATGCCCATCGGTCGCTC 76
QY	289 aagaccgagatgcagagctcacacctgaccag-ctgcccgctatcgcaatgtgttga 347
Db	77 AAGACCCGGATGACAGAGCTACAGCCTGACCCAGACCCGCCGCTATCGGAACGTGTGGA 136
QY	348 ggcctctggaggattataaagcaggcgccctatgagggcccatgagggg-gctgaacg 406
Db	137 GGCTCTCTGGAGATCATGAGACAGAGGGCCTGTGGAGGCCCATGCGGGGTGCTGAACG 196
QY	407 tcacagcaacaggcgagggcc-tgccacgcccct-ttatcttgcctgctacgaaaagt 464
Db	197 TCACAGCAACAGCGCGGGCGGTGCCACGCCCTCGTATTTTTCCTGCTACGAAAGTT 256
QY	465 aaaaagaca-ttgagtgatgtaataccacctgggggccaatagccatattgccaatggtg 523
Db	257 AAAAAAGACAGTTGAGTGACGTAATCCACCAGGGGCAATAGCCATATTGCCAATGGTG 316
QY	524 cggccggtgtgtgcaacattacttatgatcagccatgaacccctgcggaagt-ggtc 582
Db	317 CAGCCGGATGTGTGGCGACATTTACTATGATGAGCCATGAATCCAGAGTCGCTGTC 376
QY	583 aagcagaggatgcagatgtacaactcaccatccacccgggtgacagactgtgtacgggca 642

```

|||||
377 AAGCAGAGGATGAGATGTAACTCGCTCCGTACCACCGCTGACAGACTGTGTTCGGCA 436
QY 643 gttgtgcaaatgaaggccgggcttttacccagctacac-caccagctgacct 701
Db 437 GTGTGGCAAAATGAAGGGCCGGGGCTTTTACCGCAGCTACACTGACCCAGCTGACCAT 496
QY 702 gaactgtcccttccaaagccattcacttcagctatgaattctcctcagagcaacttaa 761
Db 497 GAATGTCCCTTCCAAGCCATTCACTTCATGACCTATGAGTTCTCTGCAAGAGCACATTAA 556
QY 762 cccccagagacggttacaaacccaagctcccccactctcctctctgagcttgcagcagagctgt 821
Db 557 ACCCCAGAGACGGGTACAAGCCAGCTCCCATGCTCTGTGTGGGCGCTG-GCAGGAGCAGT 615
QY 822 agtgcagcagcacaccccccaactgagctgtttgcaaaacactgctcaaacccagg--ag 879
Db 616 AGCTGCCGCCGCCGACAAAGCCCACTGGACGTTTGTAAACACACTGTTCACAGCCAGGAAT 675
QY 880 tccttgctttgaactcacacattacagagcatatcac-ggcatgcttagtgccttca- 937
Db 676 CCCTGGCTTTGAAACTCAAACTTACTTGGACACATCACAGGGCATTCCTAGTGCCTTCAG 735
QY 938 ggaactgataaagtagtggtgggtgaccgcctatttccaggggtgagccagagtaa 997
Db 736 GGACGGGTCTATCAAGTAGCGGGTTGACTGCTTACTTCCGAGGGGTGACGGGAGAGTCA 795
QY 998 ttaccagatccctcc-acagccatcgatgctgtgtatgagtcttcaaatc 1053
Db 796 TTTACCAGATCCCTCCCAACAGGCATCGCATGGTCTTGTGTATGAATTTCAAATAC 852

RESULT 9
BI103329 817 bp mRNA linear EST 26-JUN-2001
LOCUS 60288838F1 NCI_CGAP_Kid14 Mus musculus cdna clone IMAGE:5043992
DEFINITION 5', mRNA sequence.
ACCESSION BI103329
VERSION BI103329.1 GI:14554222
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 817)
COMMENT NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1120 row: 1 column: 09
High quality sequence stop: 777.
Location/Qualifiers
1. 817
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Kid14"
/lab_host="PH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. "
147 a 222 c 316 g 132 t

BASE COUNT
ORIGIN
```

```

Query Match 48.8%; Score 534.6; DB 10; Length 817;
Best Local Similarity 90.7%; Pred. No. 1.4e-96;
Matches 603; Conservative 0; Mismatches 59; Indels 3; Gaps 3;

QY 29 gttgtgccccggggccggcagggcccgccggggagccccggggagctcggcgctgtg 88
Db 153 GTGTGGGAGTGACCATGCTGTGGGCCCGGGCGAGCCCGGGAGTCTCGCGCTGTGG 212
QY 89 acgggtggttcgaagcggggcgtggccggccggcggcggcggcggcggcggcggcggc 148
Db 213 ACGGGTGGCTGCAAGCGGGGGTGGCGCGGGGGCGGGCGGGAGCGGGGGCTATC 272
QY 149 gcccccggttacacaagatccggaactcccgccggcggcggcggcggcggcggcggc 208
Db 273 AGCCCCCTGTACGGCTGGATCCGGAGTCCGGCCCGGGAATAGGAAGCCTGCCGGCTGGAG 332
QY 209 ccactgtcaaccacgcacatggtggcaggcccggtggcaggatccttgagcactgcgtga 268
Db 333 CCATGTTCACACACACATGTTGGCGGGCGCGGTGGCAGGATCCTTGAGCATTTGCGTGA 392
QY 269 tgtaccccatcgactgcgtcaagaccggcggatgaagctctacacgtgacccagctgcc 328
Db 393 TGTACCCGATCGACTGCGTCAAGACCCGGATGCGAGCGCTACAGCCTGACCCAGCCGCC 452
QY 329 gctatcgcaatgtgttgaggccctctggaggattataagaacggaggcctatggaggc 388
Db 453 GCTATCGGAACGTGTGGAGGCTCTCTGGAGATCATGAGGACAGAGGGCTGTGGAGGC 512
QY 389 ccatgagggggcgtgaacgtcacagcaacaggcggcggcggcggcggcggcggcggcggc 448
Db 513 CCATGCGGGGCTGAACTGTCACAGCAACAGCGGGGGCGCTGCCCGCCTCTATTTTG 572
QY 449 cctgctcagaaaagttaaaagacattgagtgatgtaataccaccctggggggaatagcc 508
Db 573 CCTGCTACGAAAAGTTAAAAAGACATTTAGTGCAGTAATCCACCCAGGGGCAATAGCC 632
QY 509 atattgccaatgg-tgcgcggggtgtgtgcaacattac-ttcattgatcagccatgaa 566
Db 633 ATATTGCCAATGGTTGCAGCGGATGTGGCGACATTTCTTCTATGATGAGGCAATGAA 692
QY 567 ccctgcgcgaagtgtcaagcagaggtgacagatgtacaactcaccaccacgggtgac 626
Db 693 TCCAGCGGAAGTCGTCAAGCAGAGGATGCAGATGTACAACCTGCGCGTACCACCGCGTAC 752
QY 627 agactgttcacggcagtg-tggcaaaatgaaggcggcggcggcggcggcggcggcggc 685
Db 753 AGACTGTGTTCGGGCGAGTGTGGCAAAATGACGGGGCGGGGCTTTTAAACCGAGCTACA 812
QY 686 ccacc 690
Db 813 CGAAC 817

RESULT 10
LOCUS BI057218/c
DEFINITION QV3-GN0338-230201-558-a01 GN0338 Homo sapiens cdna, mRNA sequence.
ACCESSION BI057218
VERSION BI057218.1 GI:14464748
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 576)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
```


	Query Match	47.2%	Score 517;	DB 10;	Length 927;
	Best Local Similarity	92.4%;	Pred. No. 4.5e-93;		
	Matches 544;	Conservative 0;	Mismatches 45;	Indels 0;	Gaps 0;
QY	507	ccatattgccaatgtgcggcgggtgtgtggaacattacttcattgatgcagccatgaa	566		
Db	10	CGATTATTCCTAATGGTGCAGCCGGATGTGTGGCGACATTACTTCATGATGCAGCCATGAA	69		
QY	567	ccctgcgggaagtgtcaagcagagagatgcagatgtacaaatcaccatcaccacgggtgac	626		
Db	70	TCCAGCGGAAAGTCGTCAAGCAGAGGATGCAGATGTACAACTGCGCGTACCAACCGCGTGAC	129		
QY	627	agactgtgtacgggcagttgtggcaaaatgaaggggcggggcccttttacgcgcagtacac	686		
Db	130	AGACTGTGTTTCGGGCAGTGTGGCAAAATGAAGGGGCGGGGCCCTTTTACCGCAGCTACAC	189		
QY	687	caccacgctgaccatgaacgttctctccaagccattcacttcattgacctatgaattcct	746		
Db	190	GACCCAGCTGACCATGAATGTCCCTTTCCAAAGCCATTCACTTCATGACCTATGAGTTTCCT	249		
QY	747	gcaggagcactttaacccccagagacggtgaacccccaaagctccccagctctctctgtgagc	806		

Db 250 GCAAGAGCACTTTAAACCCAGAGAGCGGTACAACCCAGCTCCCATGTGCTCTGTGGGC 309
QY ttgcagaggaactgaactccgagcagacacacccactgagcttgcacaaactgct 866
|||||
Db 310 CTGGCAGGAGCAGTAGTGGCGCGCCACACACCCACTGGACCTTTGTAAACACACTGCT 369
QY caacacccaggagtccttggctttgaactcacacattacagagacatatcacaggcatggc 926
|||||
Db 370 CAACACCCAGGAATCCCTGGCTTTGAACCTCAAAACATTACTGGACACATCACAGGCATGGC 429
QY tagtgccttcaggacggtatatacaagtagtggtgggtgacccgctatttccgaggggtgca 986
Db 430 TAGTGCCTTCAGGACGCTCTATCAAGTAGCGGGGTGACTGCTTACTTCCGAGGGGTGCA 489
QY ggcagagtaattaccagatccctccacagccatcgcatggtctgtatagttctt 1046
|||||
Db 490 GGCAGAGTCAATTTACCAAGATCCCTCCACAGCCATCGATGGTCTGTGTAAGAATTTT 549
QY caatacctaataactaaaaggcgaagagtagtgagggtggcgaagtga 1095
Db 550 CAATACTTAATCAAAAGCGGCAAGAGGAATGGAGGCGCAAGTGA 598

RESULT 12
LOCUS BI185580
DEFINITION UNL-P-FN-cy-f-02-0-UNL.s1 UNL-P-FN Sus scrofa CDNA clone
ACCESSION BI185580
VERSION BI185580.1 GI:14659989
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 574)
AUTHORS Caetano,A.R., Johnson,R.K. and Pomp,D.
TITLE Generation and sequence characterization of a normalized CDNA
JOURNAL library from swine ovarian follicles
COMMENT Unpublished (2001)
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
Oligo-dr track not found, Not I site shown in beginning of sequence
is likely internal to the message.
Seq primer: M13 -29
POLYA:No.

FEATURES
source Location/Qualifiers
1..574
/organism="Sus scrofa"
/strain="University of Nebraska, Lincoln Swine Selection
Lines"
/db_xref="taxon:9823"
/clone="UNL-P-FN-cy-f-02-0-UNL"
/clone_lib="UNL-P-FN"
/dev_stage="ADULT"
/lab_host="DHI0B (Life Technologies)"
/note="Vector: pT7p3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; The UNL-P-FN
library is a normalized library representing porcine
ovarian follicles, ranging between 2.0 to 10.0 mm in
diameter, collected during 7 days of the follicular phase
of the pig estrous cycle. This library was derived from
the library UNL-P-F2. The tag is a string of 5-6
nucleotides present between the Not I site and the
oligo-dr track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996.
TAG_SEQ=None found"

BASE COUNT 132 a 171 c 165 g 104 t 2 others
ORIGIN
Query Match 46.4%; Score 507.6; DB 10; Length 574;
Best Local Similarity 95.3%; Pred. No. 3.1e-91;
Matches 544; Conservative 0; Mismatches 25; Indels 2; Gaps 2;
QY 169 ccgaactccgcccggactacgagcgctgcgg-ctgagaccactgtcaccaagcacat 227
Db 5 CCGCTCTCCGCGCCAGACTAGCAGGCGCTGCCGNGCTGGAGCCACTGTGCACAGCACAT 64
QY 228 ggtggcagggccctggcagggatccctgagcactgcgtgatgtaccccatcgactgcgt 287
Db 65 GGTGGCGGCGCCGTGGCAGGAATCCTGGAGCACTGCGTGATGTACCCCATCGACTGGT 124
QY 288 caagaccggatgacagagctacagcctgacccagctgcccgcgtatcgcaatgtgtgga 347
Db 125 CAAGACCCGGATGCAGAGCCCTGCAGCCTGATCCCGCCGCCGCTATCGCAATGTGTGGA 184
QY 348 ggcctctgagagattataaagcagggcgctatgaagcccatgagggcgctgaacgt 407
Db 185 GGCCTCTGGAGGATTATAAGACGGAGGCGCTGTGGAGGCCCATGCGGGGGCTGAACGT 244
QY 408 cacagcaacagcgagggcgctgccacgcctttatttgcctgtctacgaaaagttaa 467
Db 245 CACAGCAACAGCGCAGGGCCTGCCAGGCCCTCTATTTCCTGCTATGAAAAGTTAA 304
QY 468 aaagacattgagtgatgaatccaccctgggggcaatagccatattgccaatgtgagcc 527
Db 305 AAAGACATTGAGTGATGATTAATCCACCTCGGGGCAATAGCCATATGTGCAATGTGCGGC 364
QY 528 cgggtgtggcaacttacttcattgatgagccataaacctcggaagtgtcaagca 587
Db 365 CGGTGGCT-ANAACATTACTTCATGATGAGCAGCCATGATCCAGAGAGTGTGCAAGCA 423
QY 588 gagatgcagatgtacaactcaccatcaccacgggtgacagactgtgtacggcagtggtg 647
Db 424 GAGGATGCAGATGTACAACCTCACCTACCACGGGTGACAGACTGTGTACGGCAGTGTG 483
QY 648 gcaaatgaagggcgccgggctttaccgcagctacacccaccagctgacacatgaacgt 707
Db 484 GCAAAACGAAGGGCGCGGGCGCTTTTACCAGCAGGTACACCACCAGCTCACCATGAACGT 543
QY 708 tcctttcaagcattcatttcattgacctat 738
Db 544 TCCCTTCCAAAGCCATCCACTTCATGACCTAT 574

RESULT 13

LOCUS BG697119
DEFINITION 602660340F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4803681 5',
mRNA sequence.
ACCESSION BG697119
VERSION BG697119.1 GI:13962975
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 908)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 560)

REFERENCE
AUTHORS
Melton D., Brown J., Kenty G., Permutt A., Lee C., Kaestner K., Lemishka I., Secorce M., Brestelli J., Gradwohl G., Clifton S., Hillier L., Marra M., Pape D., Wyllie T., Martin J., Blisain A., Schmitt A., Theising B., Ritter E., Ronko I., Bennett J., Cardenas M., Gibbons M., McCann R., Cole R., Tsagareishvili R., Williams T., Jackson Y. and Bowers Y.

TITLE
Endocrine Pancreas Consortium
Unpublished (2000)

JOURNAL
COMMENT
Other_ESTS: ic32e02.xl
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biochem.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown (brownjefas.harvard.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 430.

FEATURES
source
Location/Qualifiers
1..560
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1"
/sex="Both for embryonic & newborn, male for adult and adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Not 1; Site_2: Sal I; Five libraries representing E10.5/12.5 pancreatic bud, E16.5 pancreas, newborn pancreas, adult pancreas, and adult islets of Langerhans were separately constructed using SuperScript plasmid Library kit (Life Technologies). cDNA was made by oligo-dT priming and size-selected by column fractionation. Libraries were amplified once on solid support and plasmid DNA from each library was prepared and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an EcoT of 6. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT
134 a 158 c 163 g 105 t

ORIGIN

Query Match 45.9%; Score 502.4; DB 10; Length 560;
Best Local Similarity 93.6%; Pred. No. 3.4e-90;
Matches 524; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 215 tcaccacgacacatggtggcaggcgccgctggcaggactcctggaggactgcggtatgcc 274
Db 1 TCACCACGCACATGGTGGCGGGCGCGCTGGCAGGGATCTCTGGAGCATTTGGTGATGTACC 60

QY 275 ccacacactgcgtcaagaccgccgatcacagctctacagctgaccagctgcccgctatc 334
Db 61 CGATCGACTCGGTCAAGACCCGGATCGAGAGCTTACAGCCTGACCCAGCCGCCCGCTATC 120

QY 335 gcaatgtgttgaggccctctgaggattataagaacgaggagccctatgaggcccatga 394

```
Db 121 GGAACGTGTTGGAGGCTCTCTCGAGAATCATGAGGACAGAGGCGCTGTGGAGGCCCATGC 180
QY 395 ggggctaaactcagcaacagcgagcgctgcacacccctttatttgcctgt 454
Db 181 GGGGCTGAACGTACAGTAACAGCGCGGGGCGCTGCCACCCCTCTATTTGGCTGCT 240
QY 455 acgaaaagttaaaaaagacattgagtgtgtatgaattccacccctgggggcaatagccatttg 514
Db 241 ACGAAAAGTTAAAAAGACATTGAGTGACGTAATCACCCAGGGGCAATAGCCATATTG 300
QY 515 ccaattggtggtggcggtgtgtgtggaacattacttcaatgatcagccatgaacccctgcgg 574
Db 301 CCAATGGTGACCGGATGTGTGGGACATTACTTCATGATGACGCCATGAATCCAGCGG 360
QY 575 aagtgtcaagcagaggtgcagatgtacaactcaccattaccacccggtgacagactgtg 634
Db 361 AAGTCGTCAAGCAGAGGTGACAGTGTACAACTCGCGGTACCACCGCTGACAGACTGTG 420
QY 635 tacggcgagtgtggaatgaagggcggtggcgctttaccgcagctcacacacccagc 694
Db 421 TTCGGGCACTGTGGCAAAATGAAGGGCGGGCGCTTTTACCGCAGCTACACGACCCAGC 480
QY 695 tgaccatgaacttcttccaaagccattcaacttcaactgaacttatgaattcctgcaggagc 754
Db 481 TGACCATGAATGTCCCTTCCAAAGCCATTCACTTTCATGACCTATGAGTTTCTTCAAGAGC 540
QY 755 actttaacccccagagagcg 774
Db 541 ACTTTAAACCCCGAGAGCGG 560

RESULT 15
LOCUS BE913718 598 bp mRNA linear EST 29-SEP-2000
DEFINITION 601669072f1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5',
mRNA sequence.
ACCESSION BE913718
VERSION BE913718.1 GI:104111618
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9146 row: o column: 17
High quality sequence stop: 593.
Location/Qualifiers
1..598
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969256"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 141 a 175 c 172 g 110 t
```

ORIGIN

```
Query Match 45.9%; Score 502.2; DB 10; Length 598;
Best Local Similarity 92.3%; Pred. No. 3.8e-90;
Matches 551; Conservative 0; Mismatches 43; Indels 3; Gaps 2;

QY 195 gctgcccgtgagccactgtcacacgcacatggtggcagggcgccgtggcagggatcct 254
Db 1 GCTGCGGGCTGGAGCCACTGTACACACGCACATGGTGGCGGGCGCGTGGCAGGGATCCT 60
QY 255 ggaacactcgtgtagtatacccccatcactgcgtcgaagaccccgatgcagagtctacagcc 314
Db 61 GGAGCATTTGCGTATGTACCCGATCGCTGCTCAAGACCCCGATGCAGAGCCTACAGCC 120
QY 315 tgacccagctgcccgctatcgcaatgtgtggaagccctctgagagattataagaagga 374
Db 121 TGACCCAGCGCGCGCTATCGGAACGTGTGGAGGCTCTCTGGAGAAATCATGAGGACAGA 180
QY 375 gggcctatgagggcccatgaggggctgaacgtcacagcaacagggcgagggcctgcccc 434
Db 181 GGGCCTGTGGAGGCCCATGGGGGCTGAACGTTCACAGCAACAGGCGCGGGGCTGCCCA 240
QY 435 cgccctttatttgcctgctacgaaaagttaaaaaagacattgagtgtatgaattccaccc 494
Db 241 CGCCCTCTATTTTGCCTGTCTACGAAAAGTTAAAAAGATTAAGTGAAGTGAAGTGAAG 300
QY 495 tgggggcaatagccatattgccaatggtggcggtgtgtgtggaacattacttcatga 554
Db 301 AGGGGCAATAGCCATATTGCCAATGGTGACGCCGATGTGTGGCGACATATTCTTCATGA 360
QY 555 tgcagccatgaaccctgcggaagtgtcaagcagagagatgcagatgtacaactcacata 614
Db 361 TGCAGCCATGAATCCAGCGGAGTCTGTCAGCAGAGAGGATGCAGATGTACAACTCGCGTA 420
QY 615 ccacggggtgacagactgtgtacggcgagtggtggcaaaa--tgaaagggcgccggccttt 672
Db 421 CCACCGCGTGACAGACTGTGTTCCGGCAGTGTGGCAACACTGACAGGGGCGGGGCTTT 480
QY 673 tacc-gcagctacacacccagctgaccatgaacgttcttccagccattcacttcat 731
Db 481 TACCTGCAGCTACACGACCCAGCTGACCATGAATGTCCCTTCCAAAGCCATTCACTTCAT 540
QY 732 gacctatgaattcctgcagagcactttaacccccagagacggtacaccccaagctc 788
Db 541 GACCTATGAGTTCTCTGCAAGAGCAGCTTTAAACCCCGAGAGAGCGGTACAAACCCAGTCCC 597
```

Search completed: August 27, 2002, 01:58:09

Job time: 5374 sec
